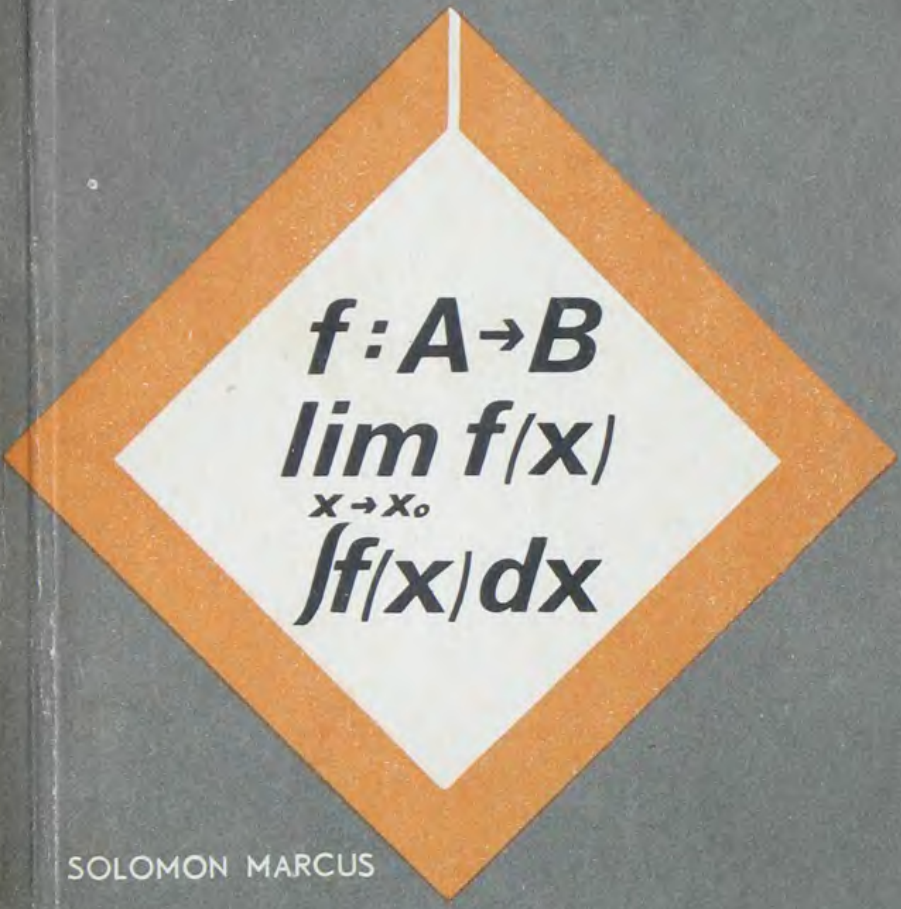
M

MATEMATICĂ

ENCICLOPEDIA DE BUZUNAR



SOLOMON MARCUS

Noțiuni de analiză matematică

SOLOMON MARCUS

Noțiuni de analiză matematică

originea, evoluția și semnificația lor

EDITURA ŞTIINŢIFICĂ • BUCUREŞTI 1967

COPERTA COLECȚIEI: T. MIRONESCU

Către cititori,

Analiza matematică face parte de multă vreme din programa analitică a numeroase institute de învățămînt superior. Pături din ce în ce mai largi de intelectuali vin în contact cu primele elemente ale acestei discipline, dar, din motive obiective, nu au posibilitatea să pătrundă mai adînc în problematica și metodele ei. Acestei largi categorii de cititori, cărora nici timpul, nici natura ocupațiilor nu le permite studierea sistematică și aprofundată a analizei matematice, dar care doresc totuși să capete o oarecare perspectivă asupra ei, li se adresează cartea de față. Ne adresăm oricui a trecut măcar o dată printr-o carte de analiză matematică de nivel universitar. Absorbit de dificultățile tehnice pe care le întîmpină, cititorul unui manual sau tratat de analiză matematică pierde uneori din vedere tocmai partea vie a dezvoltării analizei, tendințele care o animă, problemele din care se nasc noile noțiuni, sintezele pe care ea le operează. Dar o dată învinse dificultățile tehnice cu caracter local, cititorul poate simți nevoia unei retrospective în care atenția principală să fie îndreptată asupra originii, semnificației și evoluției noțiunilor de bază, încercînd să deslușească dezvoltarea ulterioară a acestora. Am fi foarte bucuroși dacă, pe acest drum, paginile care urmează îi vor fi de un oarecare ajutor.

În redactarea acestor pagini am fost însuflețiți de perspectiva largă și pasiunea pe care ni le-a transmis, de-a lungul anilor, profesorul nostru de analiză matematică, acad. Miron Nicolescu, ale cărui lecții și îndru-

mări contribuie în mod esențial, de vreo douăzeci de ani încoace, la dezvoltarea școlii românești de analiză.

Cititorii care doresc să cunoască și mai amănunțit chestiunile expuse în această carte se pot adresa manualelor și tratatelor de analiză apărute în ultima vreme la noi în țară. Dintre acestea, relevăm: pentru studenți, *Manual de analiză matematică*, în două volume, de acad. Miron Nicolescu, prof. N. Dinculeanu și prof. S. Marcus, Editura Didactică și Pedagogică, București, 1963 și 1964, și *Funcții reale și elemente de topologie generală* de acad. Miron Nicolescu, Editura Didactică și Pedagogică, București, 1967; pentru avansați și specialiști, tratatul *Analiză matematică*, în trei volume, al acad. prof. Miron Nicolescu, Editura Tehnică, București, în anii 1956, 1958, 1961.

Unele capitole ale cărții de față apar, poate, într-o lumină mai semnificativă abia după lectura unor capitole ulterioare; aceasta este, de exemplu, situația capitolelor I și III, a căror semnificație se întregește după lectura capitolului IV.

În definitivarea textului acestei cărți am profitat de observațiile prețioase ale prof. I. Barbălat și prof. C. Foiaș. Le exprim aici recunoștința mea.

SOLOMON MARCUS

7 octombrie 1966

I

De la funcții euleriene la funcții arbitrare; De la funcții arbitrare la funcții calculabile

EVOLUȚIA NOȚIUNII DE FUNCȚIE, PÎNĂ LA EULER

În mod explicit, nici grecii, nici egiptenii nu au întrebuințat noțiunea de funcție. Ei au folosit însă expresia ariei unui cerc ca funcție de diametru și expresia sumei unei progresii aritmetice ca funcție de primul termen, de rație și de numărul termenilor. Încă înainte de era noastră s-au folosit corespondențe între mărimi, definite prin tabele. Primul tabel de sinusuri trigonometrice a fost calculat în secolul al V-lea î.e.n., fiind necesar în reperarea astrelor. Cu un secol înainte de era noastră, Hiparc dă un tabel cu calculul coardelor pentru unele arce circulare. Indienii au fost și ei în posesia acestei concepții geometrice și practice. Lucrările indienilor și grecilor s-au răspîndit în Europa abia în secolul al XV-lea. La sfîrșitul secolului al XV-lea s-a observat că corespondența $n \to a^n$ stabilește o legătură între adunarea a doi întregi și înmulțirea a două puteri cu exponent întreg. Această observație a condus — la jumătatea secolului al XVI-lea — la definiția logaritmilor (Stiffel, 1544). Tot în secolul al XVI-lea, un elev al lui Copernic, Rhaeticus, a alcătuit un tabel de linii trigonometrice ale arcelor din zece în zece secunde, valorile liniilor fiind calculate cu 15 cifre exacte. Matematicienii acestui timp aveau, fără să o exprime, o idee foarte netă despre ceea ce avea să se numească mai tîrziu continuitatea funcțiilor trigonometrice. La începutul secolului al XVII-lea (1614), Neper alcătuiește primul tabel de logaritmi. Paralel cu aceasta, cercetările matematicienilor italieni și cele ale lui Descartes asupra ecuațiilor algebrice conduc, în secolele XVI și XVII, la considerarea funcțiilor algebrice.

În timpul lui Descartes (1596—1650) și Fermat (1601—1665) s-au considerat funcții algebrice definite prin curbe geometrice simple, polinoame, fracții raționale, funcții trigonometrice și inversele lor, funcția logaritmică și cea exponențială. De acest moment se leagă primul studiu sistematic al noțiunii de funcție.

În terminologia lui Newton (1642—1727) funcția este o fluentă, adică o cantitate care se scurge cu timpul; derivata e numită fluxiune, ea servește pentru a studia variația fluentei.

Cu Leibniz (1646—1716) și J. Bernoulli (1667—1748), noțiunea de funcție se eliberează de considerații accesorii și ia o formă analitică. Termenul de funcție este introdus abia la sfîrșitul secolului al XVII-lea de către J. Bernoulli.

Ținînd seama de caracterul elementar al noțiunii de funcție, este de mirare că pînă la sfîrșitul secolului al XVII-lea nu se introdusese nici un simbol special pentru a desemna o funcție. J. Bernoulli a introdus pentru prima oară un astfel de simbol, notînd cu litera X o funcție de o singură variabilă. Întrebuințarea unui simbol special pentru a desemna o necunoscută pare astfel a fi precedat cu peste 3 000 de ani întrebuințarea unui simbol pentru o funcție. Aceasta poate să pară curios unui spirit matematic modern, pentru care distanța care separă cele două întrebuințări e neînsemnată. Totuși, un școlar din zilele noastre învață încă de la 11 ani să noteze o necunoscută cu x, dar abia cu cîțiva ani mai tîrziu aude de funcții și învață să le noteze cu o literă specială.

La începutul secolului al XVIII-lea, definiția noțiunii de funcție capătă prin J. Bernoulli următoarea formă: o funcție de o mărime variabilă este o cantitate compusă într-un anumit fel din această mărime variabilă și din constante. Definiția capătă un caracter ceva mai precis prin Euler: o funcție este o expresie analitică compusă într-un fel oarecare cu ajutorul variabilelor și cu constante. Euler (1707—1783) introduce variabila complexă și funcțiile analitice, clasifică funcțiile în algebrice și transcendente și este primul care consideră funcții implicite.

STUDIUL SERIILOR TRIGONOMETRICE GRĂBEȘTE LĂRGIREA NOȚIUNII DE FUNCȚIE

Studiul vibrației coardelor în instrumentele de muzică îl conduce pe Euler să considere funcții mai generale decît cele analitice. La 1747, D'Alembert (1717—1783) a arătat că soluția problemei coardelor vibrante depinde de două funcții arbitrare supuse anumitor restricții. Euler a determinat soluția care corespunde unor anumite condiții inițiale, date de curbe trasate într-un mod arbitrar și nu numai de curbele analitice ale geometriei lui Descartes. Acestor funcții arbitrare, Euler le-a aplicat metodele rezervate, înainte de el, funcțiilor date prin expresii analitice, fără a controla dacă într-adevăr aceste funcții arbitrare admit o reprezentare analitică. În 1753, D. Bernoulli (1700—1782) dă soluția problemei coardelor vibrante cu ajutorul dezvoltării în serie trigonometrică. Compararea metodei lui cu cea a lui Euler pune problena reprezentării printr-o serie trigonometrică a unei funcții definite printr-o curbă arbitrară. În orice caz, studiul ecuației coardei vibrante a arătat de la început că anumite funcții discontinue pot fi reprezentate cu ajutorul unei serii trigonometrice convergente. Ideea de serie Fourier, adică de serie trigonometrică în care coeficienții sînt dați de formulele lui Fourier (1768—1830), aduce un spirit cu totul nou. Matematicienii se aflau în fața problemei: în ce condiții seria Fourier asociată unei funcții date este convergentă? Pentru a se răspunde la o astfel de întrebare nu mai erau suficiente metodele de calcul folosite, de pildă, de Lagrange (1736—1813) și Euler. Era nevoie, după cum spune Dirichlet, de înlocuirea calculului cu idei. Era nevoie de o definiție a corespondenței funcționale, independentă de orice formă de expresie analitică. O astfel de definiție avea să fie dată de Cauchy (1789—1857) și, mai tîrziu, de Riemann (1826—1866). Se spune că y este funcție de x, dacă fiecărei valori bine determinate a lui x îi corespunde o valoare bine determinată a lui y, indiferent de forma relației care leagă pe x de y.

Această definiție a căpătat ulterior o formă mai precisă. O funcție este dată prin două mulțimi X și Y și o regulă f care asociază fiecărui element din X un element

bine determinat din Y. Deci funcția nu este dată doar prin regula f, ci prin sistemul $<X, Y, f>$.

Pentru stabilirea în toată generalitatea și cu toată rigoarea a proprietăților relative la funcții, trebuia deci să nu mai intervină în raționamente forma particulară a expresiei algebrice sau transcendente, prin care e actualizată o funcție. Prin aceasta se generalizează însăși noțiunea de curbă. Noțiunea de funcție e eliberată de recurgerea exclusivă la intuiția geometrică. Apare definiția riguroasă a continuității, dată de Cauchy. Weierstrass construiește o funcție continuă care nu e derivabilă în nici un punct și astfel iese la iveală discrepanța dintre ideea de continuitate și suportul ei intuitiv, geometric. Și așa, spiritul de rigoare introdus de Cauchy introduce și riscurile unei noi crize. Suportul intuiției geometrice nu mai putea constitui baza raționamentelor analizei. Din această criză, analiza avea să fie scoasă de teoria mulțimilor.

Din 1807 (data primelor lucrări ale lui Fourier asupra seriilor trigonometrice) pînă la 1873, cînd apar primele lucrări ale lui Cantor (1845—1918) două direcții aveau să conducă la teoria mulțimilor. Una trece prin opera lui Cauchy, Dirichlet, (1805—1859), Fourier, Riemann, și se referă la problema integrării funcțiilor; cealaltă prin operele lui Bolzano, (1781—1848), Dedekind (1831—1916), Weierstrass (1815—1897), Cantor și se referă la principiile analizei și la delimitarea domeniului ei.

CARACTERUL ÎNȘELĂTOR AL NOȚIUNII GENERALE DE FUNCȚIE

La apariția ei, noțiunea de funcție continuă, în forma riguroasă dată de Cauchy, a indus în eroare pe matematicieni. Nu se sesiza cît de departe se poate situa această noțiune față de suportul ei intuitiv. Abia ulterior și, în unele privințe, abia în secolul nostru s-a putut înțelege întreaga generalitate a acestei noțiuni.

Cu noțiunea modernă de funcție, lucrurile s-au întîmplat invers. Atunci cînd ea a fost introdusă, matematicienii aveau impresia că se află în fața unei noțiuni foarte generale

și, tocmai de aceea, foarte săracă; în fața unei noțiuni care numai printr-o particularizare sau alta poate sa devină interesantă, să capete un conținut mai bogat, să conducă la o clasă de obiecte de oarecare însemnătate. Între expresiile analitice care făceau obiectul concepției euleriene asupra funcțiilor, pe de o parte, și funcția reala arbitrară de variabilă reală, pe de altă parte, se credea că este o distanță extraordinar de mare și nu se întrevedea vreo posibilitate ca matematica să stabilească puncte de contact între acești doi poli opuși. Această impresie înșelătoare provenea din faptul că matematicienii interpretau funcția arbitrară în sensul că sîntem liberi să alegem pentru fiecare punct al domeniului de definiție a funcției o valoare care ne convine, această alegere fiind independentă de alegerile făcute pentru celelalte puncte. Așa se explică stupoarea cu care a fost primită, la începutul acestui secol, teorema lui W. H. Young, care stabilea că pentru orice funcție de variabilă reală, în fiecare punct, exceptînd cel mult o mulțime numărabilă de puncte, mulțimea valorilor limită la stînga coincide cu mulțimea valorilor limită la dreapta. Un număr λ este o valoare limită la stînga (respectiv la dreapta) a funcției f în punctul x_0 dacă există un șir $\{x_n\}$ tinzînd către x_0, $x_n < x_0$ (respectiv $x_n > x_0$), pentru care $\lim_{n\to\infty} f(x_n) = \lambda$.

Astăzi însă care este situația? Studiul proprietăților generale ale funcțiilor a luat o asemenea amploare, încît lista proprietăților care aparțin acestor funcții a devenit impresionantă. Caracterul paradoxal pe care-l prezintă la prima vedere aceste proprietăți se șterge la o cercetare mai apropiată.

În fapt, noi nu putem decît să grupăm punctele domeniului de definiție într-un număr finit de mulțimi disjuncte și să ne dăm, pentru fiecare din aceste mulțimi, o anumită regulă de corespondență. Dacă totuși definim uneori o funcție folosind o infinitate de reguli, aceasta se explică prin faptul că regulile sînt deduse, printr-un număr finit de reguli, dintr-o colecție de asemenea finită de reguli. Această situație face ca valorile luate de funcție în diferite puncte repartizate în aceeași mulțime să nu fie complet independente, mărimea influenței unora

asupra celorlalte depinzînd de natura regulii de corespondență adoptate ca definiție a funcției pe mulțimea considerată. Deci limitele naturale ale capacității noastre de intuiție și gîndire impun noțiunii generale de funcție importante restricții.

Mai există însă și un alt aspect. Noțiunea generală de funcție constă, așa cum am mai arătat, dintr-un ansamblu de trei obiecte: o mulțime A (mulțimea de definiție), o mulțime B și o regulă f care asociază fiecărui element din A un element unic din B. O funcție realmente arbitrară ar fi aceea pentru care nu numai că aplicația f nu este supusă nici unei restricții, dar mulțimile A și B sînt și ele arbitrare, deci la rîndul lor nesupuse nici unei restricții. Însă ceea ce numim, de obicei, în analiză proprietăți ale funcțiilor arbitrare sînt, de fapt, proprietăți ale unor funcții pentru care doar regula f este arbitrară, în timp ce mulțimile A și B sînt destul de particulare. În aceste condiții, proprietățile în cauză sînt, în ultimă instanță, proprietăți ale mulțimilor A și B. Să ne referim, de exemplu, la următoarele două teoreme care se demonstrează în teoria funcțiilor reale:

a) orice funcție reală de o variabilă reală este suma a două funcții cu proprietatea lui Darboux (1842—1917) și limita unui șir de funcții cu proprietatea lui Darboux [f are proprietatea lui Darboux dacă, pentru orice interval I, $f(I)$ este un interval];

b) orice funcție reală de o variabilă reală coincide aproape peste tot cu o funcție cu proprietatea lui Darboux.

Observăm că aici mulțimile A și B sînt chiar mulțimea R a numerelor reale. Dacă A și B ar fi mulțimi arbitrare, astfel de expresii ca „proprietatea lui Darboux" și „aproape peste tot" și-ar pierde sensul, și propozițiile a) și b) n-ar mai putea fi nici măcar enunțate. Rezultă că aceste propoziții reflectă ceva din structura dreptei numerice, deoarece ele se referă la aplicațiile dreptei numerice în ea însăși.

În această ordine de idei, trebuie să citățm aici rezultatele a doi matematicieni români care, în deceniul al treilea al secolului nostru, au obținut unele rezultate interesante privind funcțiile reale arbitrare, de o variabilă reală.

Florin Vasilescu (1897—1958) în teza sa de doctorat, susținută și publicată la Paris în 1925, a întreprins un studiu sistematic al funcțiilor multiforme de variabile reale, deci al funcțiilor în definiția cărora condiția de univocitate a corespondenței este înlăturată. Ulterior, într-un articol publicat în 1927, Florin Vasilescu a ilustrat eleganța metodelor sale chiar pe exemplul funcțiilor reale obișnuite, de o variabilă reală, demonstrînd următoarea teoremă : Fiind dată o funcție reală f de o variabilă reală, există un șir $\{f_n\}$ de funcții continue, care converge către f în orice punct în care aceasta din urmă este continuă.

Alexandru Froda, în teza sa de doctorat, susținută și publicată la Paris în 1929, consacră un amplu studiu așa-numitelor proprietăți de vecinătate (adică proprietăților cu caracter local) ale funcțiilor, demonstrînd, între altele, că punctele de discontinuitate de prima specie ale unei funcții reale arbitrare, de o variabilă reală, formează o mulțime cel mult numărabilă. Pînă atunci se știa doar că dacă o funcție reală de o variabilă reală nu admite puncte de discontinuitate de a doua specie, atunci punctele ei de discontinuitate formează o mulțime cel mult numărabilă, adică o mulțime care se află în corespondență biunivocă cu mulțimea numerelor naturale sau cu o parte a acesteia. (Amintim că un punct de discontinuitate x_0 al unei funcții f este de prima specie, dacă f admite în punctul x_0 atît limită la stînga cît și limită la dreapta; în cazul contrar, punctul de discontinuitate este de a doua specie.) Ulterior, numeroase alte articole ale lui Alexandru Froda au fost dedicate unor proprietăți aparținînd tuturor funcțiilor reale de o variabilă reală.

Autorul acestei cărți a consacrat de asemenea un număr de articole funcțiilor reale arbitrare de o variabilă reală. Citățm aici unul dintre rezultatele obținute: Orice funcție reală de o variabilă reală este suprapunerea a două funcții derivabile aproape peste tot. (O proprietate are loc „aproape peste tot", dacă are loc în fiecare punct, exceptînd eventual punctele unei mulțimi de măsură nulă. O mulțime de puncte pe dreaptă este de măsură nulă, dacă ea poate fi

acoperită cu un șir de intervale de lungime totală oricît de mică.)

La sfîrșitul secolului precedent și începutul secolului nostru, cînd teoria funcțiilor reale abia își croia drum, exista un contrast între unitatea organică, eleganța și armonia pe care le prezenta teoria funcțiilor de variabilă complexă, pe de o parte, și lipsa de eleganță, caracterul migălos și complicat al raționamentelor și rezultatelor din teoria funcțiilor de variabilă reală pe de altă parte. Astăzi, teoria funcțiilor reale și-a recăpătat, pe o treaptă mai înaltă, simplitatea și pregnanța pe care le avea pe timpul lui Euler. O tendință puternică în teoria funcțiilor este întoarcerea la expresia euleriană, mai precis, la polinom. Dar această întoarcere nu se face pe vechiul drum, ci trece obligatoriu prin teoria funcțiilor generale (arbitrare).

PUNCTUL DE VEDERE STATISTIC

Într-o conferință ținută la Congresul matematicienilor români din 1929 la Cluj, acad. O. Onicescu, amintind o apropiere făcută de W. H. Young între teoria funcțiilor și calculul probabilităților, unde sub aspectul dezordonat al elementelor se poate ascunde o lege generală, observa tendința tot mai accentuată de a se descoperi proprietăți care au loc în afara unei anumite mulțimi neglijabile și le numea „proprietăți statistice ale funcțiilor". Tocmai abordarea acestui punct de vedere statistic explică bogăția și fecunditatea rezultatelor în teoria modernă a funcțiilor.

Un exemplu sugestiv în această privință îl constituie teorema lui Denjoy-Young-Saks asupra numerelor derivate ale unei funcții arbitrare. Se știe că oricărei funcții reale de o variabilă reală i se asociază, în fiecare punct, patru numere derivate, definite ca limite extreme unilaterale ale raportului dintre creșterea funcției și creșterea variabilei. (De exemplu, numărul derivat superior la dreapta al funcției $f(x)$ în punctul x_0 este, prin definiție, limita superioară a raportului $\frac{f(x) - f(x_0)}{x - x_0}$, cînd x tinde că-

tre x_0 dinspre dreapta.) Apriori, între aceste patru numere derivate sînt posibile zeci de situații și orice încercare de a le impune, în fiecare punct, anumite egalități sau inegalități atrage o restrîngere considerabilă a clasei inițiale de funcții. Cercetările profunde ale lui Denjoy, Young și Saks au arătat însă că toată această diversitate a situațiilor numerelor derivate se petrece pe o mulțime care, în multe probleme, poate fi privită ca un deșeu, anume pe o mulțime de măsură nulă. (Amintim că o mulțime de puncte pe dreaptă este de măsură nulă dacă, oricare ar fi $\varepsilon > 0$, există o familie cel mult numărabilă de intervale care acoperă mulțimea, astfel încît lungimea totală a acestor intervale este inferioară lui ε.) În afara acestei mulțimi „parazite", situația numerelor derivate ale unei funcții arbitrare poate fi caracterizată printr-o alternativă deosebit de pregnantă. Astfel, dacă numim „opuse" numărul derivat superior la stînga (respectiv la dreapta) și numărul derivat inferior la dreapta (respectiv la stînga), și „asociate" cele două numere derivate la stînga și cele două numere derivate la dreapta, se poate afirma că în fiecare punct, exceptînd punctele unei anumite mulțimi de măsură nulă, două numere derivate opuse sînt sau finite și egale sau infinite și de semne diferite; două numere derivate asociate sînt sau egale și finite, sau inegale, cel puțin unul dintre ele fiind infinit. Dar teorema are un caracter esențial statistic; fiind dat un punct oarecare, nu avem, în general, posibilitatea de a decide dacă acest punct aparține sau nu deșeului, mulțimii excepționale.

Tocmai acest punct de vedere statistic a permis stabilirea unei legături profunde între funcții generale și polinoame. O teoremă celebră a lui Weierstrass stabilește pentru orice funcție $f(x)$, continuă pe un interval închis și mărginit, existența unui șir de polinoame care converge uniform către $f(x)$. Metodele analizei clasice nu au putut, însă, arăta în ce măsură această posibilitate de aproximare prin polinoame se păstrează pentru funcții mai generale decît cele continue. Metodele subtile ale teoriei descriptive a funcțiilor, inaugurate de Baire (1874—1932) au permis să se stabilească între altele că, sacrificînd uniformitatea convergenței, posibilitatea de aproximare

prin polinoame se păstrează pentru clase întinse de funcţii discontinue, întîlnite în practica matematică. Astfel, orice funcţie derivată, orice funcţie semicontinuă şi orice funcţie cu variaţie mărginită sînt limite convergente de polinoame. (O funcţie f este superior semicontinuă în punctul x_0 dacă, fiind dat $\varepsilon > 0$, există $\eta > 0$, astfel încît $f(x) < f(x_0) + \varepsilon$ de îndată ce $|x - x_0| < \eta$. Înlocuind inegalitatea $f(x) < f(x_0) + \varepsilon$ prin $f(x) > f(x_0) + \varepsilon$, obţinem definiţia semicontinuităţii inferioare. Funcţia f este semicontinuă, dacă este în fiecare punct superior semicontinuă sau în fiecare punct inferior semicontinuă.) O astfel de teoremă are însă o valoare redusă din punct de vedere aplicativ, din cauza absenţei uniformităţii convergenţei.

În continuare, problema care se pune este de a stabili dacă nu cumva uniformitatea convergenţei reapare de îndată ce înlăturăm o anumită mulţime de puncte, mulţime care, bineînţeles, nu trebuie să fie „prea mare", adică să aibă o măsură inferioară unui număr pozitiv dat (pentru noţiunea de măsură, a se vedea capitolul IV).

În felul acesta, problema aproximării prin polinoame este pusă tocmai din punctul de vedere statistic despre care s-a vorbit mai sus. Pe de o parte, se obţine o lărgire considerabilă a clasei de funcţii care pot fi privite ca limite de polinoame; pe de altă parte, cu sacrificiul unei mulţimi „de măsură mică", este asigurată uniformitatea convergenţei şirului de polinoame către funcţia respectivă. Rezultatele acestor cercetări [întreprinse în primul rînd de Luzin (1883—1950), Borel (1871—1956) şi Fréchet] se pot rezuma astfel: fiind dată o funcţie reală $f(x)$, măsurabilă şi finită aproape peste tot pe $[a, b]$, există o mulţime E de măsură nulă, conţinută în $[a, b]$, şi un şir $P_n(x)$ de polinoame, astfel încît, pentru orice $x \in [a, b] - E$, şirul $P_n(x)$ converge către $f(x)$. Fiind dat un număr pozitiv ε, există o mulţime K, închisă şi mărginită, conţinută în $[a, b]$, astfel încît măsura (în sensul lui Lebesgue) a mulţimii $[a, b] - K$ este inferioară lui ε, iar şirul $P_1(x)$, $P_2(x), \ldots, P_n(x), \ldots$ converge, pe K, uniform către $f(x)$. (Măsura mulţimii $[a, b] - K$ este egală cu suma unei serii care are ca termeni lungimile intervalelor care nu

conțin în interior puncte din K. Evident, lungimea fiecărui astfel de interval e luată o singură dată.)

Și aici, ca și în teorema lui Denjoy-Young-Saks, este interesant să observăm caracterul esențial statistic al faptelor puse în evidență. Fiind dat un punct $x \in [a, b]$, nu sîntem în stare să decidem dacă acest punct aparține sau nu mulțimii excepționale E (sau mulțimii excepționale K).

Teorema de mai sus dă tot ce se putea aștepta în problema aproximării prin polinoame; într-adevăr, cadrul natural al teoriei funcțiilor îl furnizează astăzi clasa funcțiilor măsurabile. Nu se cunoaște nici un exemplu individual de funcție nemăsurabilă. În schimb, este remarcabil faptul că se pot da exemple individuale de funcții care scapă oricărei reprezentări analitice, adică sînt în afara clasificării lui Baire. (Funcțiile continue alcătuiesc clasa zero. Funcțiile discontinue, care sînt limite de funcții continue, alcătuiesc clasa 1; funcțiile care nu sînt continue sau de clasă 1, dar se obțin ca limite de funcții de clasa 1, se numesc funcții de clasa 2, ș.a.m.d. Această clasificare se prelungește în transfinit. O funcție care nu este de nici o clasă finită, dar este limită de funcții de clasă finită, este de prima clasă transfinită.) Nu este totuși lipsit de interes să semnalăm un rezultat al lui Sierpiński, în virtutea căruia, cu sacrificiul unei mulțimi „ceva mai mari" decît în teorema de mai sus, se poate aproxima prin polinoame orice funcție reală de variabilă reală. Iată cum se enunță acest rezultat: Fiind dată o funcție reală $f(x)$, definită pe $[a, b]$, și un număr $\varepsilon > 0$, există un polinom $P(x)$, astfel încît mulțimea punctelor $x \in [a, b]$, pentru care inegalitatea $|f(x) - P(x)| < \varepsilon$ nu are loc, este de măsură interioară inferioară lui ε. (Prin măsura interioară a unei mulțimi A conținute în $[a, b]$ se înțelege diferența dintre $b - a$ și măsura exterioară a mulțimii $[a, b] - A$. Noțiunea de măsură exterioară este definită în capitolul IV.)

FENOMENE DE CONTINUITATE LA FUNCȚII DISCONTINUE

Roadele punctului de vedere statistic în teoria modernă a funcțiilor sînt, poate, și mai vizibile atunci cînd încercăm să detectăm fenomenele de continuitate ale unei funcții

reale de variabilă reală. Astfel, pentru funcțiile din clasificarea lui Baire, I. Maximov a stabilit, în urmă cu vreo douăzeci de ani, că pentru fiecare punct x, exceptînd cel mult o mulțime numărabilă de puncte, există o mulțime perfectă P care se acumulează în x dinspre stînga și o mulțime perfectă Q care se acumulează în x dinspre dreapta, astfel încît funcția este continuă în x atît prin P, cît și prin Q. Aceasta înseamnă că pentru orice $\varepsilon > 0$ există $\eta > 0$, astfel încît din $y \in P \cup Q$. $|y - x| < \eta$ rezultă $|f(y) - f(x)| < \varepsilon$. (O mulțime P este perfectă dacă $P = P'$, unde P' este mulțimea punctelor de acumulare pentru P. Un punct x este de acumulare pentru o mulțime A dacă orice interval care conține pe x conține o infinitate de puncte din A.) În ce privește funcțiile măsurabile (a se vedea definiția lor în capitolul IV), Denjoy a stabilit, pe baza unei teoreme celebre a lui Luzin, că în fiecare punct, exceptînd cel mult o mulțime de puncte de măsură nulă, are loc așa-numita continuitate aproximativă (Hincin o numește asimptotică). Iată în ce constă acest tip de continuitate. Să notăm cu μ măsura Lebesgue (definită în capitolul IV), cu I un interval care conține în interior punctul x și cu E o mulțime măsurabilă. Să presupunem că

$$\lim_{\mu(I) \to 0} \frac{\mu(E \cap I)}{\mu(I)} = 1.$$

În aceste condiții, spunem că mulțimea E admite punctul x ca punct de densitate. Să presupunem acum că există o mulțime E care admite punctul x ca punct de densitate și astfel încît funcția f este continuă în x prin mulțimea E. Aceasta înseamnă că pentru $\varepsilon > 0$ există $\eta > 0$, astfel încît din $|y - x| < \eta$, $y \in E$, rezultă $|f(y) - f(x)| < \varepsilon$. În aceste condiții, spunem că funcția f este aproximativ continuă în punctul x. În lumina acestei definiții, teorema lui Denjoy, enunțată mai sus, are un dublu caracter statistic. Acest caracter apare atît în modul de definiție al noii specii de continuitate, definiție în care apartenența sau nonapartenența unui punct $y \neq x$ la mulțimea E nu poate fi decisă, cît și în validitatea doar în afara unei mulțimi nespecificate, dar de măsură nulă, a unei proprietăți. Oricît de artificială și de complicată ar putea să apară la prima vedere proprietatea enun-

țată, trebuie să spunem că ea are o semnificație fizică deosebită, iar noțiunea de continuitate aproximativă s-a dovedit a fi mai adecvată concepției discontinue asupra materiei decît continuitatea clasică, introdusă de Cauchy.

Dacă părăsim clasa funcțiilor măsurabile, fenomenele de continuitate devin mai puțin pregnante. Totuși, nu ne putem opri să spunem cîteva cuvinte despre o teoremă care și astăzi, la peste 40 de ani de la apariția ei, nu și-a pierdut caracterul senzațional și care provoacă uimirea oricărui matematician. Este vorba de o teoremă stabilită de H. Blumberg și care se enunță în modul următor: Fiind dată o funcție reală $f(x)$, definită pe un interval $[a, b]$, există o mulțime E conținută în $[a, b]$, densă peste tot în $[a, b]$ și astfel încît $f(x)$ este continuă, prin mulțimea E, în fiecare punct al mulțimii E. Aceasta înseamnă că fiind dat un $x \in E$ și un $\varepsilon > 0$, există un $\eta > 0$, astfel încît din $| y - x | < \eta$, $y \in E$ rezultă $| f(y) - f(x) | < \varepsilon$. Această teoremă nu a putut fi ameliorată în nici un fel. De exemplu, W. Sierpiński a arătat că există funcții pentru care mulțimea E nu poate fi aleasă în așa fel încît să fie nenumerabilă.

O CRITICĂ A NOȚIUNILOR CLASICE ȘI O PLEDOARIE PENTRU NOȚIUNILE MODERNE

Dacă din punctul de vedere al dezvoltării interne a teoriei funcțiilor diversele rafinări la care sînt supuse noțiunile fundamentale ale analizei apar explicabile, nu mai puțin interesant ar fi de văzut în ce măsură se mai păstrează acele semnificații intuitive, fizice, deosebit de izbitoare pe care le prezintă noțiunile clasice ale teoriei funcțiilor. În ce măsură concepția modernă asupra noțiunii de funcție stă față de universul fizic într-alt raport decît concepția euleriană? Iată, în această privință, opinia lui Denjoy, expusă în cadrul unei conferințe ținută la Paris în urmă cu vreo douăzeci de ani: „... Ce trebuie să așteptăm de la teoria funcțiilor de variabilă reală, în ce privește înțelegerea mai bună a realității fizice?"

„Trebuie să o privim ca stilizînd, într-un sens opus analizei clasice, aspectul materiei. Reprezentarea unui aspect din natură printr-o funcție euleriană indefinit derivabilă nu este sensibil exactă decît pentru observații la scară mare. Viteza unui avion văzut de departe poate să pară uniformă. Dacă însă din acest aparat se consideră doar o părticică și se studiază traiectoria ei în decursul unei miimi de secundă, se va identifica o vibrație a cărei viteză și amplitudine vor masca și vor domina mișcarea uniformă mijlocie a punctului material observat... O reprezentare euleriană, satisfăcătoare, la dimensiunile mijlocii, devine în întregime falsă în lumea infinitului mic".

„Teoria funcțiilor de o variabilă reală ne sugerează să asimilăm materia cu o mulțime perfectă total discontinuă. La o scară mijlocie, aceasta este echivalentă cu un număr finit de blocuri distincte și separate. O astfel de imagine poate să convină modelului, să-i dovedească armătura, așa cum funcția euleriană îi descrie aspectul exterior. Dacă însă împingem la nesfîrșit îmbucătățirea acestor blocuri, ne rămîne, la limită, un sistem de ace de recif în care nu mai găsești nici o platformă, cît de mică, pe care să pui piciorul ..."

În sensul aceleiași pledoarii pentru o concepție modernă neeuleriană asupra funcțiilor, vom da alte două citate din Denjoy (*Introduction à la thèorie des fonctions de variables réelles*, vol I, 1937, pp. 8, 9); iată de pildă, o critică a funcțiilor derivabile:

„... Sensul fizic al derivării, al diferențierii, este următorul: Cînd se observă în jurul unui punct M, într-un cîmp de dimensiune δ, un fenomen cu grosismentul $\frac{1}{\delta}$, fenomenul tinde să ia aspectul liniar pe măsură ce δ se micșorează. O suprafață fizică ia aparența unui plan, o mișcare punctuală tinde să se manifeste rectilinie și uniformă etc. Iată ce ar trebui să ne arate natura, dacă noțiunea de derivată ar fi conformă cu realitatea lucrurilor. Dar nimic nu este mai puțin adevărat. Ori de cîte ori grosismentul sub care observăm un cîmp încearcă o creștere cît de cît însemnată, aspectul fenomenului considerat este în în-

tregime răsturnat. Neregularitatea nu se va atenua, ci se va modifica mereu. Capriciul configurațiilor și al mișcărilor pare să crească atunci cînd puterea de observare crește în același timp cu micimea cîmpului.

Noțiunea de derivată este falsă din punct de vedere fizic..."

Și iată o pledoarie pentru studiul funcțiilor discontinue:

„Oriunde ai deschide ochii îți apare discontinuul. Mulțimea punctelor care pot fi unite cu un centru optic O printr-o rază rectilinie care nu întîlnește nici un obiect opac este o mulțime discontinuă. Să presupunem că o rază care pornește din punctul O este definită prin caracteristicile ei unghiulare (azimutul φ și distanța zenitală θ). Distanța r, față de O, a unui punct de pe panorama situată pe această rază este o funcție $r = f(\varphi, \theta)$, care prezintă o discontinuitate ori de cîte ori raza, presupusă mobilă, atinge conturul aparent al unui obstacol opac, după cum conturul se detașează pe un alt obiect, plasat îndărătul primului, sau, dimpotrivă, pe cer.

Funcția $r(\varphi,\theta)$ este semicontinuă inferior. Din punct de vedere fizic, aceasta este specia cea mai naturală de funcție..."

În ciuda bogatelor semnificații fizice pe care le prezintă teoria funcțiilor de punct, ideea de „funcție de punct" este, în orice caz, mai puțin adecvată fenomenului fizic decît aceea de „funcție de mulțime". Încă în urmă cu treizeci de ani, Lebesgue atrăgea atenția asupra acestei situații în cartea sa clasică asupra integrării:

„... Nu trebuie deloc să ne mire că funcțiile de domeniu se introduc și apar în fizică mult mai bine adaptate nevoilor fizicianului decît funcțiile de punct. Un punct nu e decît concepția limită a unor corpuri din ce în ce mai mici, o funcție de punct nu se poate introduce în fizică decît ca limită a unei funcții de corp, a unei funcții de domeniu. Dacă totuși se vorbește puțin despre aceste funcții, aceasta se întîmplă numai pentru că matematicienii n-au creat încă algebra și analiza funcțiilor de domeniu. În schimb, în ce privește funcțiile de punct, avem la dispoziție mijloace remarcabil de comode. Utilizînd diferite artificii, care de fapt se reduc totdeauna la a nu raționa decît pe domenii destul de speciale pentru ca

ele să nu depindă decît de un număr finit de variabile, se înlocuiește utilizarea funcțiilor de domeniu cu aceea a funcțiilor de punct...“

În anii care s-au scurs de la apariția cărții lui Lebesgue, teoria funcțiilor de mulțime a luat o mare dezvoltare. S-a edificat o teorie completă a integrării unor astfel de funcții. Totuși, teoria funcțiilor de punct este și astăzi incomparabil mai avansată. Ar fi suficient să dăm ca exemplu teoria derivării, teorie care, în ce privește funcțiile de punct, a atins întreaga perfecțiune, dar se află încă într-un stadiu relativ înapoiat în ce privește funcțiile de mulțime.

O SINTEZĂ SUPERIOARĂ: TEORIA DISTRIBUȚIILOR

A apărut însă un alt fenomen. S-a operat o sinteză a celor două tipuri de funcții — funcții de punct și funcții de mulțime — în cadrul unei teorii mai cuprinzătoare, așa-numita teorie a distribuțiilor. Este interesant că sugestia creării unei astfel de teorii provine, nemijlocit, din fizică, unde multă vreme s-a lucrat cu niște funcții avînd proprietăți imposibile din punctul de vedere al rigorii matematice, dar care își arătau, totuși, utilitatea. (De pildă, funcția lui Heaviside și funcția lui Dirac. Prima este egală cu zero pentru $x \leqslant 0$ și egală cu 1 pentru $x > 0$. Derivata funcției lui Heaviside este funcția $\delta(x)$ a lui Dirac, egală cu zero pentru $x \neq 0$ și avînd în origine o valoare atît de mare, încît

$$\int_{-\infty}^{\infty} \delta(x)\mathrm{d}x = 1.$$

În felul acesta s-a încercat — și s-a reușit — scufundarea teoriei funcțiilor de punct într-o altă teorie, mai generală, teorie în cadrul căreia proprietățile paradoxale ale funcțiilor lui Heaviside și Dirac își găsesc o explicație firească și, prin aceasta, încetează de a mai fi paradoxale. Așa cum remarca un matematician, această scufundare a teoriei funcțiilor de punct într-o teorie a unor „funcții generalizate“ se află într-o izbitoare analogie cu scufundarea teoriei numerelor raționale în teoria numerelor reale,

așa cum a fost ea efectuată, de pildă, de către Dedekind, prin celebra sa teorie a tăieturilor.

Dar cu aceasta intrăm într-o altă etapă a dezvoltării ideii de funcție, pătrundem într-un alt cerc de idei, într-un alt mod de gîndire matematică, un mod mai dinamic, în care un fenomen, pentru a putea fi înțeles, trebuie privit mereu într-o altă accepție, după natura împrejurărilor.

Noțiunea de distribuție (sau funcție generalizată) este, după cum indică chiar una dintre denumirile ei, o generalizare a noțiunii de funcție. Natura și scopul acestei generalizări amintesc întrucîtva (așa cum am semnalat și mai sus) de procesul prin care s-a trecut de la noțiunea de număr rațional la noțiunea de număr real. Pentru a obține noțiunea de număr real, se ia ca punct de plecare mulțimea numerelor raționale și se consideră toate șirurile de numere raționale care satisfac condiția lui Cauchy. (Amintim că un șir $\{a_n\}$ satisface condiția lui Cauchy dacă pentru orice număr pozitiv ε există un număr natural N, astfel încît, dacă $m > N$ și $n > N$, atunci $|a_n - a_m| < \varepsilon$.) Se introduce în mulțimea acestor șiruri o relație de echivalență și anume: două șiruri sînt echivalente dacă șirul diferență converge către zero. Clasele de echivalență astfel obținute constituie, prin definiție, numerele reale.

Pentru a obține noțiunea de distribuție, se ia ca punct de plecare mulțimea funcțiilor reale, definite pe un același interval (A, B) (unde $-\infty \leqslant A < B \leqslant \infty$). Se introduc apoi următoarele două definiții:

Despre un șir $\{F_n\}$ de funcții reale definite pe (A, B) spunem că converge pe (A, B) aproape uniform către funcția F și scriem $F_n \rightrightarrows F$, dacă F_n converge uniform către F pe orice interval compact conținut în (A, B). (De exemplu, avem $x/n \rightrightarrows 0$ pe intervalul $(-\infty, \infty)$ și $[1+x/n]^n \rightrightarrows e^x$ pe același interval.) Se demonstrează imediat că dacă un șir de funcții converge uniform pe (A, B), atunci el converge și aproape uniform pe (A, B), dar reciproca nu este adevărată. Limita unui șir de funcții continue care converge aproape uniform pe (A, B) este o funcție continuă pe (A, B). Orice șir de funcții continue care

converge aproape uniform pe (A, B) satisface condiția lui Cauchy pe (A, B). Vom nota prin $F_n \rightrightarrows$ faptul că șirul $\{F_n\}$ converge aproape uniform (către o funcție); vom nota prin $F_n \rightrightarrows\!\!\!\leftarrow G_n$ faptul că șirurile $\{F_n\}$ și $\{G_n\}$ converg aproape uniform către una și aceeași funcție.

Despre un șir $\{f_n\}$ de funcții continue pe (A, B) spunem că este fundamental pe (A, B) dacă există un șir $\{F_n\}$ de funcții și un număr întreg $K \geqslant 0$, astfel încît

$$F_n^{(k)}(x) = f_n(x) \text{ pe } (A, B) \tag{1}$$

și

$$F_n \rightrightarrows \text{ pe } (A, B). \tag{2}$$

Așa cum, pentru a se ajunge la noțiunea de număr real, se introducea o relație de echivalență în mulțimea șirurilor de numere raționale care satisfac condiția lui Cauchy, pentru a se ajunge la noțiunea de distribuție, se introduce o relație de echivalență în mulțimea șirurilor fundamentale de funcții continue pe (A, B).

Vom spune că două astfel de șiruri $\{f_n\}$ și $\{g_n\}$ sînt echivalente și vom scrie

$$\{f_n\} \sim \{g_n\},$$

dacă există două șiruri $\{F_n\}$ și $\{G_n\}$ de funcții reale definite pe (A, B) și un număr întreg $K \geqslant 0$, astfel încît derivata de ordinul K a funcției F_n (respectiv G_n) pe (A, B) să fie egală cu f_n (respectiv g_n), iar șirurile $\{F_n\}$ și $\{G_n\}$ converg pe (A, B) aproape uniform către una și aceeași funcție. Cu alte cuvinte, avem pe (A, B)

$$F_n^{(k)}(x) = f_n(x), \qquad G_n^{(k)}(x) = g_n(x), \tag{3}$$

$$F_n \rightrightarrows\!\!\!\leftarrow G_n. \tag{4}$$

Se poate arăta că valoarea K din definiția de mai sus poate fi înlocuită prin orice alt număr întreg mai mare decît K.

Fiind date două șiruri fundamentale $\{f_n\}$ și $\{g_n\}$ de funcții continue pe (A, B), ele sînt echivalente dacă și numai dacă șirul format prin alternare

$$f_1, g_1, f_2, g_2, \ldots, f_n, g_n, \ldots \tag{5}$$

este fundamental pe (A, B). Într-adevăr, dacă șirul (5) este fundamental, există un număr întreg $k \geqslant 0$ și două șiruri $\{F_n\}$ și $\{G_n\}$ astfel încît sînt satisfăcute relațiile (3), iar șirul

$$F_1, G_1, F_2, G_2, \ldots, F_n, G_n, \ldots \tag{6}$$

converge aproape uniform pe (A, B), ceea ce implică îndeplinirea condiției (4). Deci condițiile (3) și (4) sînt satisfăcute. Reciproc, dacă condițiile (3) și (4) sînt satisfăcute, atunci șirul (6) converge aproape uniform pe (A, B), deci șirul (5) satisface relațiile (1) și (2).

Deoarece am numit relația $\sim$ o relație de echivalență, este necesar să arătăm că ea este, într-adevăr, reflexivă, simetrică și tranzitivă. Verificarea reflexivității și a simetriei este destul de simplă pentru a o lăsa pe seama cititorului; vom verifica aici doar tranzitivitatea relației $\sim$. Să presupunem că avem șirurile fundamentale $\{f_n\}$, $\{g_n\}$ și $\{h_n\}$ de funcții continue pe (A, B), astfel încît $\{f_n\} \sim \{g_n\}$ și $\{g_n\} \sim \{h_n\}$. Există deci un număr întreg k nenegativ și două șiruri $\{F_n\}$ și $\{G_n\}$ satisfăcînd condițiile (3) și (4); există de asemenea un număr întreg nenegativ l și două șiruri $\{\bar{G}_n\}$ și $\{H_n\}$ satisfăcînd relații de tipul (3) și (4):

$$\bar{G}_n^{(l)}(x) = g_n(x), \quad H_n^{(l)}(x) = h_n(x), \quad \bar{G}_n \rightrightarrows\!\!\!\leftleftarrows H_n.$$

În baza unei observații de mai sus, putem presupune că $l = k$. În aceste condiții, punînd $\bar{H}_n = G_n - \bar{G}_n + H_n$, obținem

$$F_n^{(k)}(x) = f_n(x), \quad \bar{H}_n^{(k)}(x) = h_n(x), \quad F_n \rightrightarrows\!\!\!\leftleftarrows \bar{H}_n,$$

deci avem $\{f_n\} \sim \{h_n\}$ și proprietatea de tranzitivitate este stabilită. Rezultă că șirurile fundamentale de funcții continue pe (A, B) se repartizează în clase de echivalență întocmai ca și șirurile de numere raționale care satisfăceau condiția lui Cauchy. Așa cum o clasă de șiruri Cauchy echivalente constituia, prin definiție, un număr real, o clasă de șiruri fundamentale de funcții continue constituie, prin definiție, o distribuție sau o funcție generalizată. Orice șir fundamental $\{f_n\}$ de funcții continue pe (A, B) va reprezenta și funcția generalizată corespunzătoare.

Dacă $\{f_n\} \sim \{g_n\}$, atunci cele două șiruri reprezintă aceeași funcție generalizată.

Este clar că orice șir staționar (deci cu termeni egali) de funcții continue pe (A, B) este fundamental pe (A, B). Deci orice funcție continuă pe (A, B) este în același timp o funcție generalizată pe (A, B). Reciproca nu este însă adevărată; nu orice funcție generalizată se reduce la o funcție obișnuită, continuă pe (A, B). Astfel, de pildă, este funcția generalizată definită de șirul

$$f_n(x) = \sqrt{\frac{n}{2\pi}}\, e^{-nx^2/2} .$$

Se poate arăta că acest șir este fundamental pe $(-\infty, \infty)$; funcția generalizată corespunzătoare reprezintă funcția δ a lui Dirac, despre care a fost vorba la începutul acestui paragraf. Această funcție a lui Dirac este un exemplu de funcție generalizată care nu poate fi identificată cu o funcție continuă.

Se poate arăta că două funcții continue distincte definesc totdeauna două funcții generalizate, de asemenea distincte. Noțiunea de funcție generalizată este deci o extensiune a noțiunii de funcție continuă. Se poate arăta că noua noțiune înglobează și o clasă largă de funcții discontinue. Analogia dintre acest nou concept și conceptul de număr real poate fi rezumată în felul următor:

a) numere reale	a) funcții generalizate
b) numere raționale	b) funcții continue
c) șiruri Cauchy de numere raționale	c) șiruri fundamentale de funcții continue
d) șiruri Cauchy echivalente	d) șiruri fundamentale echivalente
e) numărul real este o clasă de șiruri Cauchy echivalente de numere raționale	e) funcția generalizată este o clasă de șiruri fundamentale echivalente de funcții continue
f) numere iraționale	f) funcții generalizate care nu se reduc la funcții continue

Analogia desfășurată mai sus se referă la structura logică a celor două concepte. Dar ea funcționează nu mai puțin în ceea ce privește finalitatea conceptelor considerate.

Noțiunea de număr real se introduce pentru a legitima anumite operații cu numere raționale, cum ar fi operația de extragere a rădăcinii pătrate din 2 sau aceea a logaritmării anumitor numere pozitive, operații care nu pot căpăta un sens atîta vreme cît rămînem în cadrul mulțimii numerelor raționale. Conceptul de funcție generalizată îndeplinește o finalitate asemănătoare. Astfel, pentru a da un singur exemplu, se știe că operația de derivare nu este posibilă pentru orice funcție continuă; există chiar funcții continue care nu sînt derivabile în nici un punct. În schimb, se poate arăta că operația de derivare capătă sens pentru orice funcție generalizată; orice astfel de funcție admite derivate de orice ordin. În particular, funcțiile continue vor admite și ele, în noul cadru, derivate de orice ordin, dar aceste derivate nu mai sînt funcții obișnuite, ci funcții generalizate. Derivarea dubioasă a funcției lui Heaviside (și obținerea, pe această cale, a funcției lui Dirac) devine, în lumina noului concept de derivată, o operație perfect legitimă. Iată cum se definește operația de derivare a unei funcții generalizate. Se arată mai întîi că printre șirurile fundamentale echivalente care definesc o funcție generalizată se află obligatoriu și un șir fundamental de polinoame $\{p_n\}$. Derivata de ordinul m a funcției generalizate considerate va fi, prin definiție, funcția generalizată definită de șirul $\{p_n^{(m)}\}$, adică de șirul derivatelor de ordinul m ale polinoamelor p_n (se arată că șirul acestor derivate este fundamental).

În cazul particular în care o funcție continuă admite derivate ordinare continue pînă la ordinul m, aceste derivate coincid cu derivatele în sensul teoriei funcțiilor generalizate.

Teoria funcțiilor generalizate sau — cum se spune mai frecvent — teoria distribuțiilor își are originile în fizica matematică, în teoria ecuațiilor cu derivate parțiale și în alte teorii matematice. Bazele ei au fost puse de către matematicianul francez Laurent Schwartz, după cel de-al doilea război mondial. Prezentarea făcută mai sus urmează o altă variantă, datorită matematicienilor polonezi I. Mikusiński și R. Sikorski. La noi în țară, teoria distribuțiilor a constituit obiectul multor cercetări. Remarcăm, în special, pe cele întreprinse de regretatul Alexandru Ghica, fost membru al Academiei și profesor la Universitatea din Bucu-

rești, de profesorul Gh. Marinescu, membru corespondent al Academiei (autorul monografiei *Espaces vectoriels pseudo-topologiques et théorie des distributions*, Akademie Verlag, Berlin, 1964) și, în colaborare cu prof. Romulus Cristescu, autor al monografiei *Unele aplicații ale teoriei distribuțiilor*, Editura Academiei, București, 1966, de profesorul Romulus Cristescu (autorul cărții *Elemente de analiză funcțională și teoria distribuțiilor*, Editura Tehnică, București, 1966) și de alții.

O altă generalizare a noțiunii de funcție, care coincide cîteodată cu cea de distribuție, o constituie noțiunea de operator în sensul lui Mikusiński. Și această noțiune poate fi definită printr-un proces de completare. Fie $C([0, T))$ spațiul funcțiilor continue definite pe $[0,T)$ (unde $0<T\leqslant\infty$). Fie $f, g \in C([0, T))$. Vom defini o funcție $g * f$ prin

$$(\alpha) \qquad g * f(x) = \int_0^x g(x - t) f(t)\, dt, \quad x \in [0, T).$$

Un șir $\left\{f_n\right\}_{n=1}^{\infty} \subset C([0,T))$ se spune fundamental, dacă există $g \in C([0,T))$ neidentic nulă, astfel ca $g * f_n \rightrightarrows$. Două astfel de șiruri $\{f_n\}$ și $\{f'_n\}$ se numesc echivalente, dacă există $h \in C([0, T))$ neidentic nulă, astfel încît $h * (f_n - f'_n) \rightrightarrows 0$. Se arată ușor că în familia șirurilor fundamentale relația de echivalență de mai sus constituie o egalitate algebrică; prin urmare, această familie se împarte în clase de echivalență. Un operator în sensul lui Mikusiński este tocmai o astfel de clasă. Definiția inițială a lui Mikusiński diferă de aceasta. Folosirea definiției de mai sus a devenit posibilă în urma unei frumoase teoreme de aproximare, datorită lui Ciprian Foiaș (publicată în *Studia Math*, 1961).

Să definim cîteva operații cu operatori. Astfel, dacă $\{f_n\}$ și $\{f'_n\}$ sînt două șiruri fundamentale reprezentînd operatorii φ și φ', iar g și g' sînt funcțiile care apar în definirea faptului că șirurile $\{f_n\}$ și $\{f'_n\}$ sînt fundamentale, atunci $\{f''_n\}$, unde $f''_n = g * f'_n + g' * f_n$, este de asemenea un șir fundamental (se consideră $k = g * g'$ și se arată ușor că avem $k * f''_n \rightrightarrows$). Se poate arăta că clasa definită de $\{f''_n\}$ nu depinde de reprezentanții $\{f_n\}$ și $\{f'_n\}$ aleși, așa că prin considerațiile de mai sus se definește un operator ψ care, prin

definiție, este suma $\varphi + \varphi'$ a celor doi operatori φ, φ'. În mod analog $\{f_n * f'_n\}$ este de asemenea fundamental (căci $k * f_n * f'_n = (g * f_n) * (g' * f'_n)$). Se arată de asemenea că clasa θ determinată de $\{f_n * f'_n\}$ nu depinde de reprezentanții $\{f_n\} \in \varphi$ și $\{f'_n\} \in \varphi'$ aleși. Prin definiție, acest operator θ este produsul $\varphi \times \varphi'$ al celor doi operatori. Mulțimea operatorilor înzestrați cu aceste două operații $+$ și $*$ devine un inel, numit inelul lui Mikusiński.

Orice funcție $f \in C([0, T))$ poate fi identificată cu un operator, anume cu clasa care conține șirul fundamental $\{f_n\}$, unde $f_n = f$ $(n = 1, 2, \ldots)$. Se arată ușor că operația de adunare obișnuită și cea de „înmulțire" definită de (α) pentru funcții coincid în acest caz cu cele definite pentru operatori. ■

GRADE DE EFECTIVITATE

Diferitele funcții utilizate în analiză prezintă o mare varietate în ceea ce privește gradul lor de efectivitate, caracterul mai mult sau mai puțin constructiv al regulii care asociază fiecărui element din mulțimea de definiție un element din mulțimea valorilor. Între funcțiile cele mai „neefective", cum ar fi cele definite cu ajutorul axiomei lui Zermelo (enunțată și discutată în capitolul IV), și funcțiile definite printr-o regulă care comportă doar un număr finit de pași aplicați numerelor naturale există o scară bogată de trepte intermediare, corespunzătoare funcțiilor în a căror definiție se folosesc tipuri intermediare de efectivitate, cum ar fi utilizarea totalității numerelor transfinite de a doua clasă, utilizarea procedeului diagonal al lui Cantor (procedeu ilustrat de demonstrația nenumărabilității mulțimii numerelor reale) etc. A apărut din ce în ce mai pronunțat divorțul dintre existența matematică și construcția matematică. Încă la sfîrșitul secolului trecut, o dată cu amploarea luată de studiul mulțimilor infinite și ca urmare a apariției unor paradoxuri și antinomii generate de acest studiu, au început între matematicieni discuții aprinse asupra legitimității, în matematică, a unor obiecte care nu pot fi obținute într-un mod suficient de efectiv.

Aceste discuții, care au continuat, cu intensitate crescîndă, în primele decenii ale secolului nostru, sînt acum de domeniul trecutului. În secolul trecut, matematicieni ca Kronecker (1823—1891) se declarau partizani înfocați ai finitismului, intenționau să excludă din matematică întreaga teorie a lui Cantor privind mulțimile infinite. Kronecker vedea în teoria mulțimilor infinite o construcție pur verbală. Intuiționismul matematic, reprezentat în special de Brouwer, deși mai puțin zelos decît Kronecker, respingea totuși o bună parte din teoriile lui Cantor și Zermelo (1871—1953), din logistica lui Peano (1858—1932) și Russel, din ideile lui Hilbert (1862—1943) relativ la fundamentele matematicii. Acest exclusivism a fost înlocuit, treptat, printr-o înțelegere din ce în ce mai nuanțată a existenței matematice. A devenit clar că matematica are nevoie de toate gradele de efectivitate, că noțiuni și teoreme de un grad relativ mai mic de efectivitate ajută la înțelegerea și aprofundarea unor teorii cu un grad de efectivitate mai mare. A înlătura din matematică obiectele a căror existență nu are un caracter constructiv înseamnă a limita în mod dăunător perspectiva acestei științe. Dar, în același timp, se impune o delimitare cît mai netă a acelei părți din matematică pe care o putem obține prin procedee de un anumit grad de efectivitate. În particular, a devenit din ce în ce mai stringentă necesitatea de a delimita acea parte a matematicii care poate fi introdusă în mașină. Așa s-a născut matematica constructivă și, ca un capitol important al acesteia, analiza matematică constructivă. Noțiunea de bază aici este aceea de funcție calculabilă, înțelegînd prin aceasta o funcție ale cărei valori pot fi obținute cu ajutorul unui algoritm normal. Dar ...

CE ESTE UN ALGORITM NORMAL?

Această noțiune a fost introdusă, în urmă cu vreo zece ani, de către matematicianul sovietic A.A. Markov, fiul marelui matematician A.A. Markov (1856—1922). Uneori, pentru a nu-i confunda, se vorbește despre A.A. Markov-tatăl și A.A. Markov-fiul. A.A. Markov-tatăl a fost un mare

specialist în teoria probabilităţilor şi a devenit celebru prin procesele pe care le-a introdus şi le-a studiat, procese care sînt azi cunoscute sub numele de procese Markov. În ceea ce-l priveşte pe A.A. Markov-fiul, opera sa de căpetenie este *Teoria algoritmilor*, apărută în 1954, şi despre care vom căuta să dăm unele noţiuni în cele ce urmează.

Să considerăm o mulţime A de elemente numite litere. A constituie un alfabet. Orice şir finit de litere, dintre care unele se pot repeta, se numeşte cuvînt peste alfabetul A sau, dacă nu există pericol de confuzie, pur şi simplu cuvînt. Mulţimea cuvintelor peste A va fi notată prin $\mathcal{C}(A)$. Un element din $\mathcal{C}(A)$ este deci de forma:

$$a_1 a_2 \ldots a_i \ldots a_n,$$

unde $a_i \in A$ $(i = 1, \ldots, n)$. Lungimea unui cuvînt este, prin definiţie, numărul literelor (distincte sau nu) care alcătuiesc acest cuvînt.

Fiind date două cuvinte, $\alpha = a_1 \ldots a_i \ldots a_n$ şi $\beta = b_1 \ldots b_j \ldots b_m$, vom defini cuvîntul $\alpha\beta$ ca fiind egal cu

$$a_1 a_2 \ldots a_i \ldots a_n b_1 \ldots b_j \ldots b_m.$$

În mod analog, fiind date trei cuvinte α, β, γ, putem defini cuvîntul $\alpha\beta\gamma$ format prin alăturarea literelor din α, β şi γ în ordinea în care ele figurează în aceste cuvinte.

Se consideră şi cuvîntul vid, notat prin θ şi definit prin relaţia $\alpha = \alpha\theta = \theta\alpha$, oricare ar fi cuvîntul α.

Fie două cuvinte α şi β. Vom spune că β este inclus în α, dacă există două cuvinte γ şi δ, astfel încît

$$\alpha = \gamma\beta\delta. \tag{7}$$

Nu se exclude posibilitatea ca unul dintre cuvintele γ, δ sau fiecare dintre ele să fie cuvîntul vid. Rezultă deci că fiecare cuvînt este inclus în el însuşi, iar cuvîntul vid este inclus în orice cuvînt.

Dacă β este inclus în α, alegem, dintre diferitele reprezentări ale lui α, sub forma (7), pe aceea în care γ are lungimea cea mai mică. (Se poate demonstra că există o astfel de reprezentare şi numai una.) Reprezentarea lui α sub forma (7), cu γ de lungime minimă, se numeşte prima incluziune a lui β în α.

Fie acum $\alpha_1, \ldots, \alpha_n$ și $\beta_1, \ldots, \beta_n$ cuvinte din $\mathcal{C}(B)$, unde B este un alfabet care conține alfabetul A.

Să considerăm schema

(S) $\alpha_1 \to \beta_1$, $\alpha_2 \to \beta_2$, ..., $\alpha_i \to \beta_i$. ..., $\alpha_j \to \beta_j$, ..., $\alpha_n \to \beta_n$.

Cu ajutorul schemei (S) vom defini o funcție $\mathcal{A}$ care asociază unora dintre cuvintele din $\mathcal{C}(B)$ cîte un cuvînt bine determinat din $\mathcal{C}(B)$.

Fie π un cuvînt din $\mathcal{C}(B)$. Dacă nici unul dintre cuvintele $\alpha_1, \alpha_2, \ldots, \alpha_n$ nu este inclus în π, atunci punem prin definiție $\mathcal{A}(\pi) = \pi$. Dacă există un α_j inclus în π, atunci, alegînd pe j astfel încît, pentru nici o valoare $k<j$, α_k să nu fie inclus în π, considerăm prima incluziune a lui α_j în π: $\pi = \gamma\alpha_j\delta$. După cum indică schema (S), înlocuim pe α_j cu β_j și obținem cuvîntul $\pi_1 = \gamma\beta_j\delta$. Dacă săgeata de rang j ar fi fost urmată de un punct (așa cum se întîmplă cu săgeata de rang i), atunci am fi pus $\mathcal{A}(\pi) = \pi_1$. În cazul contrar, cuvîntul π_1 este supus aceluiași tratament ca și cuvîntul π. Dacă nici un cuvînt α_i ($1 \leqslant i \leqslant n$) nu este inclus în π_1, atunci punem $\mathcal{A}(\pi) = \pi_1$. În cazul contrar, fie α_k inclus în π_1, astfel încît α_m să nu fie inclus în π_1 pentru nici un $m<k$. Fie $\pi_1 = \varepsilon\alpha_k\omega$ prima incluziune a lui α_k în π_1. Urmînd indicația săgeții de rang k din schema (S), înlocuim pe α_k prin β_k și obținem cuvîntul $\pi_2 = \varepsilon\beta_k\omega$. Dacă săgeata de rang k este urmată de un punct, atunci punem $\mathcal{A}(\pi) = \pi_2$. În cazul contrar, aplicăm lui π_2 același tratament pe care l-am aplicat anterior lui π și lui π_1 ș.a.m.d.

Dacă, procedînd așa cum am arătat mai sus, nu ajungem niciodată la o săgeată urmată de un punct sau la un cuvînt π_i în care să nu fie inclus nici unul dintre cuvintele $\alpha_1, \ldots \alpha_n$, atunci funcția $\mathcal{A}$ nu se definește pentru cuvîntul π, cu alte cuvinte nu are sens să vorbim despre $\mathcal{A}(\pi)$. Dacă însă ajungem la o săgeată urmată de un punct sau la un cuvînt π_i, în care nu este inclus nici unul dintre cuvintele $\alpha_1, \ldots, \alpha_n$, atunci cuvîntul $\mathcal{A}(\pi)$ există și este unic determinat.

Funcția $\mathcal{A}$, definită mai sus, este, prin definiție, un algoritm normal peste alfabetul B. Schema (S) este schema algoritmului $\mathcal{A}$. Un algoritm normal este determinat prin schema (S).

Vom ilustra noțiunea de algoritm normal pe exemplul operației de adunare aplicată numerelor naturale. Cu acest prilej, se va înțelege și sensul considerării unui alfabet B mai bogat decît alfabetul inițial A.

Fie un alfabet A format dintr-o singură literă, pe care o vom reprezenta prin cifra 1. Vom reprezenta numerele naturale în baza 1; astfel, numărul 7 se reprezintă prin

$$1\ 1\ 1\ 1\ 1\ 1\ 1\ .$$

Să considerăm P numere naturale $N_1, N_2, \ldots, N_p$. Fiecare dintre numerele $N_1, \ldots, N_p$ se poate reprezenta ca un cuvînt peste alfabetul A, ceea ce revine la reprezentarea sa în baza 1; pentru a reprezenta însă sistemul tuturor acestor P numere naturale, mai este nevoie de un simbol; vom folosi în rolul acestui simbol auxiliar o steluță: $*$. În felul acesta, obținem un alfabet B care conține alfabetul A: $B = \{1, *\}$.

Sistemul celor P numere naturale se reprezintă, în alfabetul considerat, prin cuvîntul

$$\alpha = \underbrace{1\ 1\ldots1}_{N_1 \text{ ori}} * \underbrace{1\ 1\ldots1}_{N_2 \text{ ori}} * \ldots * \underbrace{1\ 1\ldots1}_{N_p \text{ ori}}. \tag{8}$$

Trebuie să găsim un algoritm $\mathcal{A}$ care transformă cuvîntul α în cuvîntul

$$\mathcal{A}(\alpha) = \underbrace{1\ 1\ldots\ldots\ldots\ldots1}_{N_1 + N_2 + \ldots + N_p \text{ ori}},$$

cuvînt care constituie reprezentarea în baza 1 a sumei $N_1 + \ldots. + N_p$.

Să considerăm schema (S) următoare:

$$\to \theta.$$

(Amintim că θ este cuvîntul vid.) Vom arăta că funcția $\mathcal{A}$ definită de această schemă este tocmai algoritmul normal căutat.

Fie un cuvînt β din $\mathcal{C}(A)$, deci care nu conține cuvîntul $*$. Rezultă că $\mathcal{A}(\beta) = \beta$. Aceasta este natural, pentru că revine la situația în care cele p numere care trebuie adunate se reduc la unul singur; acest număr va fi, evident, rezultatul adunării.

Fie acum un cuvînt α care conține atît litera 1, cît și litera $*$. Deci, α este de forma (8). Prima incluziune a lui $*$ în α este

$$\alpha = \gamma * \delta,$$

unde

$$\gamma = \underbrace{1 \ldots 1}_{N_1 \text{ ori}}, \quad \delta = \underbrace{1 \ldots 1}_{N_2 \text{ ori}} * \ldots * \underbrace{1 \ldots 1}_{N_p \text{ ori}}.$$

Aplicînd schema (S), obţinem cuvîntul

$$\alpha_1 = \gamma\theta\delta = \gamma\,\delta = \underbrace{1 \ldots\ldots\ldots 1}_{N_1 + N_2 \text{ ori}} * \underbrace{1 \ldots 1}_{N_3 \text{ ori}} * \ldots * \underbrace{1 \ldots 1}_{N_p \text{ ori}}.$$

Prima incluziune a lui $*$ în α_1 este

$$\alpha_1 = \gamma_1 * \delta_1,$$

unde

$$\gamma_1 = \underbrace{1 \ldots\ldots\ldots 1}_{N_1 + N_2 \text{ ori}}, \quad \delta_1 = \underbrace{1 \ldots . 1}_{N_3 \text{ ori}} * \ldots * \underbrace{1 \ldots\ldots 1}_{N_p \text{ ori}}.$$

Aplicînd schema (S), obţinem cuvîntul

$$\alpha_2 = \gamma_1\,\theta\,\delta_1 = \gamma_1\delta_1 = \underbrace{1 \ldots\ldots\ldots 1}_{N_1 + N_2 + N_3 \text{ ori}} * \underbrace{1 \ldots . 1}_{N_4 \text{ ori}} * \ldots * \underbrace{1 \ldots . 1}_{N_p \text{ ori}}.$$

Continuînd în acest fel, ajungem, după un număr finit de paşi, la cuvîntul

$$\alpha_{p-1} = \underbrace{1 \ldots\ldots\ldots 1}_{N_1 + \ldots + N_p \text{ ori}}$$

şi, deoarece $*$ nu este inclus în α_{p-1}, rezultă că

$$\mathcal{A}(\alpha) = \alpha_{p-1}.$$

Să considerăm un algoritm normal $\mathcal{A}$ definit de schema (S). Pentru a fixa ideile, presupunem că singura săgeată urmată de un punct („săgeată finală") este săgeata de indice j. Introducem pentru fiecare număr natural i, $1 \leqslant i \leqslant n$ operatorul O_i, care constă în a înlocui, în prima incluziune a cuvîntului α_i în cuvîntul prelucrat, pe α_i prin β_i. Mai introducem, pentru fiecare număr natural i, $1 \leqslant i \leqslant n$, condiţia logică C_i = cuvîntul α_i este inclus în cuvîntul prelucrat. În sfîrşit, mai introducem condiţia logică identic falsă, condiţie pe care o reprezentăm prin C_ω.

Să considerăm acum următorul șir de operatori și de condiții logice:

$$\underset{n+1}{\downarrow} C_1 \cdots \underset{n+j-1}{\downarrow} \underset{n+j+1}{\downarrow} \underset{2n}{\downarrow} C_1 \overset{1}{\uparrow} O_1 C_\omega \overset{n+1}{\underset{1}{\uparrow}} \downarrow C_2 \overset{2}{\uparrow} O_2 C_\omega \overset{n+2}{\uparrow} \cdots$$

$$\cdots \underset{j-1}{\downarrow} C_j \overset{1}{\uparrow} O_j C_\omega \overset{n+j}{\uparrow} \cdots \underset{n-1}{\downarrow} C_n \overset{n}{\uparrow} O_n C_\omega \overset{2n}{\uparrow} \underset{n}{\downarrow} \underset{n+j}{\downarrow}.$$

Această schemă prelucrează fiecare cuvînt α în cuvîntul $\mathcal{A}(\alpha)$, unde $\mathcal{A}$ este algoritmul normal definit de schema (S). Dacă C_1 se verifică, cu alte cuvinte dacă α_1 este inclus în cuvîntul α, atunci se aplică operatorul O_1, situat imediat la dreapta lui C_1, deci se înlocuiește α_1 prin β_1 în prima incluziune a lui α_1 în α. Urmează apoi condiția C_ω care, fiind identic falsă, nu se verifică; deci, urmînd indicația săgeților de indice $n + 1$, se începe din nou prin a se verifica condiția C_1 etc. Dacă C_1 nu se verifică, cu alte cuvinte dacă α_1 nu este inclus în α, atunci, urmînd indicația săgeților de indice 1, se verifică condiția C_2. Dacă α_2 este inclus în α, atunci C_2 este îndeplinită și se aplică operatorul O_2, situat imediat la dreapta; urmează apoi verificarea condiției C_ω care, fiind identic falsă, nu e satisfăcută; deci, urmînd săgețile de indice $n + 2$, se reîncepe cu verificarea condiției C_1 etc.

O situație specială o prezintă C_j și O_j. Dacă C_j se verifică, atunci se aplică operatorul O_j, apoi urmează condiția identic falsă C_ω; deci, urmînd săgețile de indici $n + j$, ajungem la extremitatea din dreapta a schemei, prin urmare $\mathcal{A}(\alpha)$ a fost găsit.

FUNCȚIILE CALCULABILE, PUNCT DE PLECARE ÎN ANALIZA CONSTRUCTIVĂ

Cu ajutorul noțiunii de funcție calculabilă (adică de funcție ale cărei valori pot fi determinate cu ajutorul unui algoritm normal) se definește noțiunea de număr real constructiv după cum urmează. Se introduce mai întîi noțiunea de șir calculabil de numere raționale, ca fiind un șir de numere raționale pentru care există o funcție calculabilă care dă termenul de rang n. Apoi se definește noțiunea de

regulator de convergență al unui șir $\{a_n\}$, ca fiind o funcție f definită pe mulțimea numerelor naturale, cu valori în aceeași mulțime, și astfel încît pentru $k \geqslant f(i)$, $l \geqslant f(i)$ avem

$$|a_k - a_l| < \frac{1}{2^i}.$$

Un șir calculabil de numere raționale este, prin definiție, calculabil convergent, dacă admite drept regulator de convergență o funcție calculabilă. Se introduce o relație de echivalență în mulțimea șirurilor calculabile de numere raționale, care sînt calculabil convergente. Clasele de echivalență astfel obținute sînt, prin definiție, numerele reale constructive.

Dezvoltarea analizei constructive urmează programul analizei clasice: convergența, numărul real, limita unei funcții, continuitatea, derivata, formulele de medie, integrala, funcțiile cu variație mărginită, teoria măsurii etc. Însă această similitudine de program nu implică o similitudine corespunzătoare în ceea ce privește rezultatele obținute. Astfel, orice funcție reală calculabilă, de o variabilă reală constructivă, este continuă (în sens constructiv). Apropiind acest rezultat de faptul că în teoria distribuțiilor orice funcție continuă este indefinit derivabilă, ne dăm seama că multe dintre punctele de vedere moderne relative la funcții introduc un proces de regularizare, de înlăturare a anomaliilor; funcțiile tind să devină „din ce în ce mai cumsecade".

FUNCȚII CALCULABILE ÎN SENSUL LUI TURING

În 1936, deci înainte de construirea primelor calculatoare electronice, A.M. Turing a imaginat un model matematic de mașină care a permis introducerea unui concept riguros de calculabilitate. Calculabilitatea în sensul lui Markov este echivalentă cu cea în sensul lui Turing, dar deosebirile de ordin metodologic care există între ele le conferă grade diferite de comoditate în problemele atît de variate în care apare conceptul de calculabilitate. Deoa-

rece mașina Turing prezintă o deosebită importanță în logica matematică și în teoria mașinilor matematice, o vom prezenta în cele ce urmează, rămînînd ca după aceasta să expunem o a treia variantă a conceptului de calculabilitate, aceea care se introduce cu ajutorul funcțiilor recursive.

Înainte de a prezenta definiția matematică a mașinii Turing, vom căuta să dăm o idee intuitivă despre acest concept, imaginîndu-l realizat ca o mașină fizică. O mașină Turing se compune din : *o unitate centrală*, care este susceptibilă de un număr finit de stări; *o bandă infinită*, divizată în celule, în fiecare celulă putînd fi înscris un simbol și numai unul (inițial, datele care urmează să fie prelucrate de către mașină sînt înscrise pe această bandă și tot pe ea sînt înscrise și rezultatele intermediare); *un organ de citire și scriere*, care asigură comunicarea între unitatea centrală și banda infinită. Acest organ nu poate acționa simultan în mai multe celule ale benzii, el putînd executa exclusiv următoarele operații: înlocuirea unui simbol citit prin altul; deplasarea benzii, cu o celulă, la stînga sau la dreapta (fig. 1. I)

Inițial, mașina se află într-o stare determinată (numită *starea inițială*) și citește un simbol înscris într-o celulă a benzii. Ca rezultat al acestei activități, mașina trece, eventual, într-o altă stare a unității centrale, ceea ce are ca urmare fie înlocuirea simbolului citit prin altul, fie suprimarea lui, fie deplasarea benzii, cu o celulă, la stînga sau

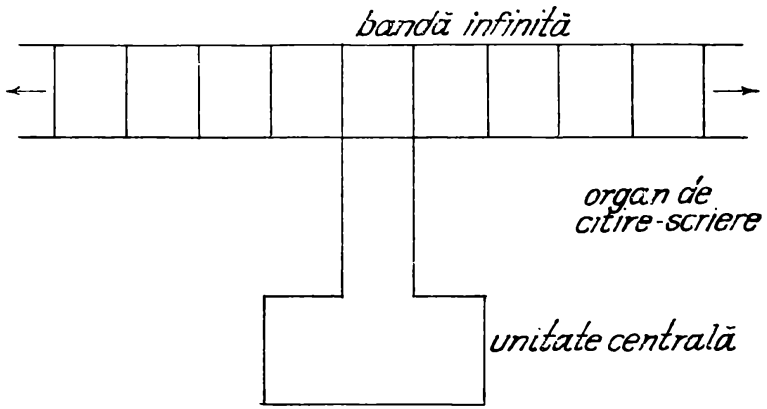


Fig. 1.I.

la dreapta. Să observăm că suprimarea unui simbol este un caz particular de înlocuire a simbolului, fiind o înlocuire cu așa-numitul *simbol vid*, a cărui prezență într-o celulă marchează tocmai faptul că celula respectivă este liberă. După ce au fost prelucrate toate simbolurile secvenței înscrise pe bandă, secvența obținută este rezultatul prelucrării primei secvențe cu ajutorul mașinii. Se poate însă întîmpla ca mașina să nu fie în stare să prelucreze unele secvențe, în sensul că activitatea mașinii se perpetuează, nefurnizînd prelucrarea completă a secvenței după nici un număr finit de etape (oricît de mare ar fi acest număr).

Să trecem acum la definirea riguroasă a conceptului de mașină Turing. Menționăm că varianta pe care o adoptăm nu coincide cu varianta dată de Turing, fiind o ameliorare a acesteia, datorită lui Emil L. Post și a altor matematicieni, variantă care corespunde la ceea ce Martin Davis numește *mașină Turing simplă.*

Fie un alfabet finit $(\mathcal{S}) = \{S_0, S_1, \cdots, S_n\}$, numit *alfabetul mașinii.* S_0 este numit uneori *blancul*, fiind interpretat ca simbolul vid și notat de multe ori prin litera B. A înlocui pe S_i cu S_0 înseamnă deci pur și simplu a șterge pe S_i.

Fie $Q = \{q_1, q_2, \cdots, q_t\}$ o mulțime finită de elemente numite *stări*; q_1 este așa-numita *stare inițială.*

Fie D și S două simboluri noi, avînd semnificația de „deplasare la dreapta" și, respectiv, „deplasare la stînga". Numim *expresie* orice șir finit de elemente din mulțimea $\mathcal{S} \cup Q \cup \{D, S\}$ (un același element putînd apărea de mai multe ori într-un șir).

O clasă specială de expresii ne va reține atenția; aceea a cvadrupletelor. Numim *cvadruplet* orice expresie care prezintă una din următoarele trei forme:

$$q_i S_j S_k q_l$$

$$q_i S_j D q_l$$

$$q_i S_j S q_l$$

Numim *cuplu inițial* al unui cvadruplet cuplul ordonat al primilor doi termeni ai săi. De exemplu, în cvadrupletul $q_i S_j S_k q_l$ cuplul inițial este $q_i S_j$. Numim *cuplu terminal*

al unui cvadruplet cuplul ordonat al ultimilor doi termeni ai cvadrupletului.

Numim *mașină Turing* o colecție de obiecte $<\mathcal{S}, Q, D, S, \Lambda>$, unde $\mathcal{S}$, Q, D și S au semnificația de mai sus, iar Λ este o mulțime finită de cvadruplete care diferă, două cîte două, prin cuplul inițial.

Cuplul inițial al unui cvadruplet din Λ reprezintă o *situație a mașinii*, în timp ce *evoluția ulterioară a mașinii* este indicată de cuplul terminal. Condiția ca două cvadruplete din Λ să difere obligatoriu prin cuplul inițial imprimă mașinii Turing o *funcționare deterministă*, în sensul că unei anumite situații a mașinii îi corespunde o evoluție ulterioară unic determinată.

Cele trei tipuri de cvadruplete corespund la trei tipuri diferite de funcționare a mașinii. Un cvadruplet de forma $q_iS_jS_kq_l$ indică faptul că mașina, aflîndu-se în starea q_i și citind simbolul S_j, îl înlocuiește pe acesta cu S_k și trece în starea q_l. (Dacă avem $k = 0$, simbolul S_j este pur și simplu șters.) Un cvadruplet de forma $q_iS_jDq_l$ indică faptul că mașina, aflîndu-se în starea q_i și citind simbolul S_j imprimă benzii o deplasare la dreapta și trece în starea q_l. Un cvadruplet de forma $q_iS_jSq_l$ indică faptul că mașina, aflîndu-se în starea q_i și citind simbolul S_j, imprimă benzii o deplasare la stînga și trece în starea q_l. Să menționăm faptul că, în fiecare din cele trei cazuri este posibil ca $l = i$, deci ca mașina să persiste în starea în care deja se află.

Pentru a face o analogie cu un calculator fizic, am putea spune că o stare (internă) a mașinii Turing corespunde unei ipostaze a tuburilor și angrenajelor unui calculator fizic. Într-adevăr, funcția acestor tuburi și angrenaje este aceea de a determina actul următor al calculatorului. Banda unei mașini Turing corespunde memoriei unui calculator fizic. Dar aici apare o deosebire esențială între o mașină Turing și un calculator fizic; în timp ce acesta din urmă admite doar o regiune finită (memoria-esențial finită — a calculatorului) de stocare a datelor inițiale și a expresiilor intermediare, formate în cursul calculului, la o mașină Turing această regiune (care se confundă cu banda mașinii) este infinită. Mașina Turing este deci o mașină ideală.

Pentru a obține o imagine mai dinamică, de proces, a activității unei mașini Turing, vom introduce noțiunea de *descriere instantanee*, înțelegînd prin aceasta o expresie printre ai cărei termeni se află un singur termen care reprezintă o stare, termen care nu este situat la extremitatea din dreapta a expresiei; în plus, expresia nu conține nici o apariție a lui D sau a lui S. De exemplu expresia $q_2BS_3q_3S_1$ nu este o descriere instantanee, deoarece conține două apariții de stări. Nici expresiile $S_1S_2S_3q_2$ și q_1S_2D nu sînt descrieri instantanee, deoarece prima conține o stare la extremitatea ei din dreapta, iar a doua conține simbolul D. În schimb, expresiile $q_1S_1S_2S_2$ și $S_1S_0S_2S_1q_2S_1$ sînt descrieri instantanee.

O descriere instantanee α este o *descriere instantanee a mașinii* Z dacă q_i din α este o stare a mașinii Z, iar termenii din α care nu sînt stări, sînt elemente din alfabetul mașinii Z. Descrierea instantanee explică activitatea mașinii; prin suprimarea lui q_i se obține conținutul benzii la un anumit moment, iar dacă q_i este urmat de S_j, cuplul q_iS_j definește situația mașinii la momentul respectiv.

Fie α și β două descrieri instantanee asociate mașinii $Z = <\mathcal{S}, Q, D, S, \Lambda>$. Spunem că β *derivă direct* din α și scriem $\alpha \rightarrow \beta$ (sau, dacă există posibilitate de confuzie în ceea ce privește mașina considerată, $\alpha \rightarrow \beta(Z)$), dacă are loc unul din următoarele cinci cazuri (unde P și R sînt șiruri oarecare — eventual vide — de simboluri ale alfabetului $\mathcal{S}$):

1) există P și R astfel încît $\alpha = Pq_iS_jR$, $\beta = Pq_lS_kR$, unde $q_iS_jS_kq_l$ este un cvadruplet din Λ;

2) există P și R astfel încît $\alpha = Pq_iS_kR$, $\beta = PS_jq_lS_kR$, unde $q_iS_jDq_l$ este un cvadruplet din Λ;

3) există P și R astfel încît $\alpha = Pq_iS_j$, $\beta = PS_jq_lS_0$, unde $q_iS_jDq_l$ este un cvadruplet din Λ;

4) există P și R astfel încît $\alpha = PS_kq_iS_jR$, $\beta = Pq_lS_kS_jR$, unde $q_iS_jSq_l$ este un cvadruplet din Λ;

5) există R astfel încît $\alpha = q_iS_jR$, $\beta = q_lS_0S_jR$, unde $q_iS_jSq_l$ este un cvadruplet din Λ.

O descriere instantanee β este *terminală* relativ la mașina Z, dacă nu există nici o descriere instantanee γ a lui Z, astfel încît $\beta \rightarrow \gamma(Z)$.

Derivația directă are un caracter univoc, în sensul că dacă descrierile instantanee α, β și γ ale lui Z sînt astfel încît $\alpha \to \beta$ și $\alpha \to \gamma$, atunci $\beta = \gamma$. Într-adevăr, să presupunem că $\alpha = Pq_iS_jQ$, $\beta = Pq_lS_kQ$ și $\gamma = Pq_rS_mQ$ (în celelalte cazuri se procedează în mod analog). Rezultă că mașina, aflîndu-se în starea q_i și citind pe S_j, înlocuiește pe S_j prin S_k sau S_m și trece în starea q_l sau q_r. Deci, dacă am avea $k \neq m$ sau $l \neq r$, s-ar contrazice caracterul deterministic al mașinii.

Se mai poate arăta că dacă Z și Z' sînt două mașini Turing cu proprietatea că orice cvadruplet din Z este și un cvadruplet din Z' (caz în care scriem $Z \subset Z'$), atunci din faptul că $\alpha \to \beta$ relativ la Z rezultă că $\alpha \to \beta$ relativ la Z'.

Numim *calcul al mașinii* Z orice șir finit $\alpha_1, \alpha_2 \ldots \alpha_p$ de descrieri instantanee ale lui Z, astfel încît $\alpha_i \to \alpha_{i+1}$ pentru $1 \leqslant i \leqslant p-1$, iar α_p este terminală. Despre situația instantanee α_p spunem că este rezultatul prelucrării de către Z a descrierii α_1 și scriem

$$\alpha_p = \mathrm{Rez}_Z(\alpha_1).$$

Pentru a ilustra noțiunile introduse mai sus, să considerăm cîteva exemple de mașini Turing:

a) fie mașina $Z = <\mathcal{S}, Q, D, S, \Lambda>$, unde $\mathcal{S} = \{S_0, S_1, S_2\}$, $Q = \{q_1, q_2, q_3\}$, iar Λ este formată din următoarele cvadruplete:

(1) $q_1S_0Dq_1$
(2) $q_1S_2Dq_1$
(3) $q_1S_1Dq_2$
(4) $q_2S_1Sq_3$
(5) $q_2S_0Dq_1$
(6) $q_2S_2Dq_1$
(7) $q_3S_1S_2q_3$.

Să ne dăm acum un anumit cuvînt pe alfabetul ($\mathcal{S}$), fie el $S_1S_0S_2S_1S_1S_0S_1$, și să cercetăm modul în care mașina Z prelucrează acest cuvînt. Activitatea mașinii începe (ca la orice mașină Turing) în mod obligatoriu din starea inițială q_1, deci prima descriere instantanee este

$$q_1S_1S_0S_2S_1S_1S_0S_1. \qquad (I)$$

Această descriere corespunde situației q_1S_1 a mașinii Z. Observăm că există un cvadruplet al lui Z, cel numerotat cu (3), al cărui cuplu inițial este q_1S_1. Acest cvadruplet arată că mașina, aflîndu-se în starea q_1 și citind simbolul S_1, imprimă benzii o deplasare la dreapta și trece în starea q_2. Descrierea instantanee (I) va fi deci urmată de descrierea instantanee

$$S_1q_2S_0S_2S_1S_1S_0S_1, \qquad \text{(II)}$$

așa cum rezultă din cazul 2, care apare în definiția derivației directe (avem I→II). Această descriere corespunde situației q_2S_0 a mașinii, situație care apare în cvadrupletul (5), al cărui cuplu inițial este într-adevăr q_2S_0. Acest cvadruplet arată că mașina, aflîndu-se în starea q_2 și citind simbolul S_0, imprimă benzii o deplasare la dreapta și trece în starea q_1. Descrierea instantanee (II) va fi deci urmată de descrierea instantanee (rezultată tot din cazul 2 al definiției relației →)

$$S_1S_0q_1S_2S_1S_1S_0S_1. \qquad \text{(III)}$$

Această descriere corespunde situației q_1S_2 a mașinii, situație care apare în cvadrupletul (2). Acest cvadruplet arată că mașina, aflîndu-se în starea q_1 și citind simbolul S_2, imprimă benzii o deplasare la dreapta și persistă în starea S_1. Descrierea instantanee (III) va fi deci urmată de descrierea instantanee (obținută tot din cazul 2 al definiției relației →)

$$S_1S_0S_2q_1S_1S_1S_0S_1. \qquad \text{(IV)}$$

Această decriere corespunde situației q_1S_1 a mașinii, situație care apare în cvadrupletul (3). Efectul acestui cvadruplet asupra descrierii instantanee (IV) are ca rezultat descrierea instantanee (obținută tot prin referire la cazul 2 din definiția relației →)

$$S_1S_0S_2S_1q_2S_1S_0S_1. \qquad \text{(V)}$$

În această descriere apare situația q_2S_1, căreia îi corespunde cvadrupletul (4), al cărui cuplu inițial este într-adevăr q_2S_1. Acest cvadruplet arată că mașina, aflîndu-se în starea q_2 și citind simbolul S_1, imprimă benzii o deplasare

la stînga și trece în starea q_3. Aceasta are ca efect înlocuirea descrierii (V) prin descrierea

$$S_1S_0S_2q_3S_1S_1S_0S_1, \qquad \text{(VI)}$$

așa cum rezultă din cazul 4 care apare în definiția derivației directe.

Descrierea instantanee astfel obținută corespunde situației q_3S_1 a mașinii, care indică acțiunea cvadrupletului (7). Acest cvadruplet arată că mașina, aflîndu-se în starea q_3 și citind simbolul S_1, înlocuiește pe S_1 prin S_2 și persistă în starea q_3. Urmează deci descrierea instantanee

$$S_1S_0S_2q_3S_2S_1S_0S_1. \qquad \text{(VII)}$$

Aici apare situația $q_3\,S_2$, căreia nu-i corespunde în A nici un cvadruplet. Aceasta înseamnă că activitatea mașinii încetează. Șirul de descrieri instantanee (I), (II), (III), (IV), (V), (VI), (VII) este un calcul al mașinii Z, deoarece descrierea VII s-a dovedit a fi terminală și, pe de altă parte, avem $I \to II \to III \to IV \to V \to VI \to VII$. Putem deci scrie $S_1S_0S_2q_3S_2S_1S_0S_1 = \mathrm{Rez}_Z(q_1S_1S_0S_2S_1S_1S_0S_1)$, ceea ce înseamnă că rezultatul prelucrării cuvîntului $S_1S_0S_2S_1S_1S_0S_1$ de către mașina Z este cuvîntul obținut din VII prin ștergerea lui q_3: $S_1S_0S_2S_2S_1S_0S_1$.

După cum am văzut, mașina Z a putut efectua prelucrarea cuvîntului $S_1S_0S_2S_1S_1S_0S_1$ într-un număr finit de etape, în sensul că după un număr finit de etape activitatea mașinii încetează. Se poate însă întîmpla ca înscriind pe banda unei mașini Turing un anumit cuvînt, prelucrarea sa de către mașină să continue indefinit, în sensul că la fiecare etapă se obține o descriere instantanee neterminală în raport cu mașina considerată. Să luăm, de exemplu, cuvîntul $S_1S_2S_2$ și să-l înscriem pe banda mașinii analizate mai sus, la punctul a). Pornind din starea q_1 și citind primul simbol al cuvîntului considerat, se obțin, aplicînd succesiv cvadrupletele (3), (6), (2) și (1), următoarele descrieri instantanee:

$$q_1S_1S_2S_2$$
$$S_1q_2S_2S_2$$
$$S_1S_2q_1S_2$$
$$S_1S_2S_2q_1S_0$$
$$S_1S_2S_2S_0q_1S_0.$$

Observăm că prin efectul cvadrupletului (1), penultima descriere instantanee a condus la o descriere instantanee asupra căreia acționează același cvadruplet (1). (Adăugarea, la extremitatea din dreapta, a unei descrieri instantanee a simbolului vid S_0 nu afectează cu nimic descrierea respectivă.) În felul acesta, situația q_1S_0 devine perpetuă, iar aplicarea cvadrupletului (1) se repetă de o infinitate de ori, conducînd la descrieri instantanee de forma

$$S_1S_2\underbrace{S_2S_0S_0\ldots S_0}_{n \text{ ori}}q_1S_0.$$

unde n este un număr natural arbitrar;

b) să considerăm mașina $Z = <\$, Q, D, S, \Lambda>$, unde $\$ = \{S_0, S_1, S_2, S_3\}$, $Q = \{q_1, q_2, q_3\}$, iar Λ este alcătuită din următoarele cvadruplete: $q_1S_0Dq_1$, $q_1S_1Dq_1$, $q_1S_3Dq_1$, $q_1S_2S_3q_2$, $q_2S_3Sq_3$, $q_3S_0S_3q_3$, $q_3S_1S_3q_3$, $q_3S_2S_3q_3$. Un exemplu de calcul al mașinii Z este următorul $S_1q_1S_0S_3S_2 \rightarrow$ $\rightarrow S_1S_0q_1S_3S_2 \rightarrow S_1S_0S_3q_1S_2 \rightarrow S_1S_0S_3q_2S_3 \rightarrow S_1S_0q_3S_3S_3$, deci avem

$$\text{Rez}_Z\,(S_1q_1S_0S_3S_2) = S_1S_0q_3S_3S_3.$$

Rezultă că cuvîntul $S_1S_0S_3S_2$ este prelucrat în cuvîntul $S_1S_0S_3S_3$. Un alt exemplu de calcul al aceleiași mașini este următorul:

$$q_1S_2 \rightarrow q_2S_3 \rightarrow q_3S_0S_3 \rightarrow q_3S_3S_3, \text{ deci}$$

$$\text{Rez}_Z(q_1S_2) = q_3S_3S_3,$$

ceea ce înseamnă că cuvîntul S_2 este prelucrat de Z în cuvîntul S_3S_3.

Iată acum și un exemplu de descriere instantanee α pentru care $\text{Rez}_Z(\alpha)$ nu este definit. Fie $\alpha = S_1q_1S_0S_3S_1$. Avem $\alpha \rightarrow S_1S_0q_1S_3S_1 \rightarrow S_1S_0S_3q_1S_1 \rightarrow S_1S_0S_3S_1q_1S_0 \rightarrow$ $S_1S_0S_3S_1S_0q_1S_0 \rightarrow \ldots$; observăm că situația q_1S_0 se repetă într-o infinitate de descrieri instantanee succesive, fiecare derivînd direct din cea precedentă.

Pentru ca o mașină Turing să efectueze nu numai calcule *simbolice* ci și calcule *numerice*, este necesar să introducem o reprezentare convenabilă a numerelor. Alegem

un simbol de bază, de exemplu pe S_1, pe care-l notăm aici 1, și reprezentăm cu ajutorul lui orice număr natural. Vom folosi bineînțeles și simbolul vid (blancul) B. Precizăm că fiind dat un simbol k, folosim notațiiiile

$$k^n = \underbrace{kk...k}_{n \text{ ori}}, \quad k^0 = B.$$

Fiecărui număr natural n îi asociem expresia $\bar{n}$ egală, prin definiție, cu 1^{n+1}. De exemplu, $\bar{3} = 1^4 = 1111$. Fiecărui sistem ordonat de p numere naturale (0 fiind și el considerat număr natural) $<n_1, n_2, \cdots, n_p>$ îi asociem expresia

$$\overline{<n_1, n_2, \ldots n_p>} = \bar{n}_1 B \bar{n}_2 B \ldots . B \bar{n}_p.$$

De exemplu,

$$<\overline{2, 3, 0}> = 111B1111B1.$$

Notațiile introduse vor fi folosite pentru exprimarea datelor inițiale (de intrare). În ceea ce privește expresiile de ieșire, vom folosi o altă notație, și anume: fiind dat un șir finit M de elemente din $\mathcal{S} \cup Q$, notăm cu $<M>$ numărul de apariții ale lui 1 (deci ale lui S_1) în M. De exemplu: $<11BS_4q_3> = <S_1S_1BS_4q_3> = <S_1S_1S_0S_4q_3> = 2$; $<q_3q_2S_5> = 0$. Observăm că

$$\overline{<m - 1>} = <1^m> = m.$$

Este ușor de demonstrat egalitatea $<MN> = <M> + <N>$.

Fie acum o mașină Turing Z. Pentru fiecare număr întreg pozitiv n asociem lui Z o funcție $\psi_Z^{(n)}$ definită pe o parte a produsului cartezian $\mathcal{N}^n$ și cu valori în $\mathcal{N}$ (mulțimea numerelor naturale) după cum urmează. Fiind dat sistemul ordonat $(m_1, m_2, \ldots, m_n)$ de n numere naturale, punem

$$\alpha_1 = q_1\overline{(m_1, m_2, \ldots, m_n)},$$

obținem astfel o descriere instantanee a lui Z (deoarece q_1 este starea inițială a lui Z, iar $S_1 = 1$ este aici un simbol din alfabetul mașinii Z). Dacă există un calcul al lui Z care începe cu α_1, fie acest calcul $\alpha_1, \ldots, \alpha_p$, atunci funcția $\psi_Z^{(n)}$

este definită în punctul $(m_1, m_2, \ldots, m_n)$ din $\mathcal{N}^n$ și punem, prin definiție,

$$\psi_Z^{(n)}(m_1, m_2, \ldots, m_n) = <\alpha_p> = <\mathrm{Rez}_Z(\alpha_1)>.$$

Dacă nu există nici un calcul al mașinii Z care începe cu descrierea instantanee α_1, deci dacă

$$\mathrm{Rez}_Z(\alpha_1)$$

nu este definit (ceea ce revine la perpetuarea activității lui Z), atunci funcția $\psi_Z^{(n)}$ nu este definită în punctul $(m_1, m_2, \ldots, m_n)$. Cu alte cuvinte, funcția $\psi_Z^{(n)}$ este definită exact în acele puncte ale lui $\mathcal{N}^n$ care furnizează descrieri instantanee pe care mașina Z le poate prelucra într-un număr finit de etape.

Fie acum o funcție f definită pe o parte a produsului cartezian $\mathcal{N}^n$ și cu valori în $\mathcal{N}$. Vom spune că f este o *funcție parțial calculabilă în sensul lui Turing*, dacă există o mașină Turing Z, astfel încît

$$f = \psi_Z^{(n)}.$$

Aceasta înseamnă că funcțiile f și $\psi_Z^{(n)}$ au aceeași mulțime de definiție și iau valori egale în fiecare punct în care sînt definite. Se mai spune în acest caz că *mașina Z calculează funcția f*. Dacă, în plus, funcția f este *totală* (adică definită în fiecare punct al produsului cartezian $\mathcal{N}^n$), atunci spunem că f este o *funcție calculabilă în sensul lui Turing*.

Pentru ilustrare, să arătăm că funcția $f(x, y) = x + y$ definită pe $\mathcal{N} \times \mathcal{N}$ este o funcție calculabilă în sensul lui Turing. În acest scop, construim o mașină $Z = <\mathcal{S}, Q, D, S, \Lambda>$, unde $\mathcal{S} = \{B, 1\}$, iar componentele Q și Λ vor fi alese astfel încît

$$\psi_Z^{(2)}(x, y) = x + y.$$

Înscriind pe bandă reprezentările $\bar{x}$ și $\bar{y}$ separate printr-un blanc (adică prin B), vom cere mașinii să furnizeze pe $x + y$, ținînd seamă că $x + y = <\bar{x}> + <\bar{y}> - 2$. O soluție a problemei este dată de $Q = \{q_1, q_2, q_3\}$ și Λ, alcătuită din următoarele cvadruplete: $q_1 1 B q_1$, $q_1 B D q_2$, $q_2 1 D q_2$, $q_2 B D q_3$, $q_3 1 B q_3$. Primul cvadruplet are ca efect ștergerea unei unități din $\bar{x}$; următoarele trei cvadruplete au

:a efect deplasarea benzii pînă la prima unitate din $\bar{y}$, iar ıltimul cvadruplet determină ştergerea primei unităţi din $\bar{y}$ şi oprirea maşinii, deoarece situaţia q_3B nu constituie :uplul iniţial al nici unui cvadruplet. Însă în acest moment banda conţine $x + y$ unităţi. Fie deci

$$\alpha_1 = q_1(\overline{m_1,\ m_2}) = q_1\bar{m}_1B\bar{m}_2.$$

Avem: $\alpha_1 = q_1 1 1^{m_1} B\ 1 1^{m_2} \to q_1 B 1 1^{m_1} B 1 1^{m_2} \to$

$\to Bq_2\, 1^{m_1} B 1 1^{m_2} \to \cdots \to B 1^{m_1} q_2 B\ \ 1 1^{m_2} \to B 1^{m_1} Bq_3\, 1 1^{m_2} \to$

$\to B 1^{m_1} Bq_3 B 1^{m_2}$, ultima descriere instantanee fiind terminală în raport cu Z. Aşadar, $\psi_Z^{(2)}(m_1, m_2) = <\mathrm{Rez}_Z(\alpha_1)> =$

$$= <B1^{m_1}Bq_3B1^{m_2}> = m_1 + m_2.$$

Să dăm acum două exemple mai simple de funcţie calculabilă în sensul lui Turing.

Fie *funcţia succesor* $S(x) = x + 1$. Vom alege o maşină pentru care $q_1\bar{m}$ este o descriere terminală, oricare ar fi m. De exemplu, fie $Z = <\mathcal{S}, Q, D, S, \Lambda>$, unde $\mathcal{S} = \{B, 1\}$, $Q = \{q_1\}$, iar Λ este alcătuită din unicul cvadruplet q_1BBq_1. Se observă că avem $\psi_Z^{(1)}(m) = <q_1\bar{m}> = m + 1$.

Fie acum *funcţia identică* $I(x) = x$. Această funcţie poate fi calculată cu următoarea maşină Turing: $Z = <\mathcal{S}, Q, D, S, \Lambda>$, unde $\mathcal{S} = \{B, 1\}$, $Q = \{q_1\}$, iar Λ este alcătuită din unicul cvadruplet $q_1 1 Bq_1$. Într-adevăr, avem

$$q_1\bar{n} = q_1\, 1 1^n \to\ q_1 B 1^n, \quad \text{deci} \quad \psi_Z^{(1)}(n) = <q_1 B 1^n> = n.$$

Unele funcţii destul de simple necesită pentru a fi calculate, o maşină Turing destul de complicată. Aşa se întîmplă, de exemplu, cu funcţia $f(x, y) = x - y$. Această funcţie este definită doar pe o parte a lui $\mathcal{N} \times \mathcal{N}$, anume în punctele (x, y) pentru care $x \geqslant y$. Deci nu se pune problema de a arăta că funcţia este calculabilă, ci doar de a arăta că ea este

parțial calculabilă în sensul lui Turing. În construirea unei mașini corespunzătoare ne vom călăuzi după următoarea idee: înscriind pe bandă reprezentările $\bar{x}$ și $\bar{y}$, separate printr-un blanc, mașina va trebui ca, la fiecare etapă să șteargă unitatea situată la extremitatea din stînga a lui x, precum și unitatea situată la extremitatea din dreapta a lui y. Vom defini deci mașina $Z = <\mathcal{S}, Q, D, S, \Lambda>$, unde $\mathcal{S} = \{B, 1\}$, $Q = \{q_1, q_2, q_3, q_4, q_5, q_6, q_7, q_8, q_9\}$ iar Λ este alcătuită din următoarele 17 cvadruplete:

$q_1 1 B q_1$, $q_1 B D q_2$, $q_2 1 D q_2$, $q_2 B D q_3$, $q_3 1 D q_3$, $q_3 B S q_4$, $q_4 1 B q_4$, $q_4 B S q_5$, $q_5 1 S q_6$, $q_6 1 S q_6$, $q_6 B S q_7$, $q_7 1 S q_8$, $q_7 B D q_9$, $q_8 1 S q_8$, $q_8 B D q_1$, $q_9 B D q_9$, $q_9 1 S q_9$.

Pentru a arăta că $\psi_Z^{(2)}(m_1, m_2) = m_1 \dot{-} m_2$, plecăm de la descrierea instantanee $\alpha_1 = q_n(\overline{m_1, m_2}) = q_1 \overline{m}_1 B \overline{m}_2$ și, presupunînd că $m_1 \geqslant m_2$, punem $k = m_1 \dot{-} m_2$. Avem:

$$\alpha_1 = q_1\, 1^{m_1+1} B 1^{m_2+1} = q_1 1 1^{m_2} 1^k B 1^{m_2} 1 \rightarrow$$

$$\rightarrow q_1 B 1^{m_2} 1^k B_1^{m_2} 1 \rightarrow B q_2 1^{m_2} \rightarrow B q_2 1^{m_2} 1^k B 1^{m_2} 1 \rightarrow \cdots \rightarrow$$

$$\cdots \rightarrow B 1^{m_2} 1^k q_2 B 1^{m_2} 1 \rightarrow B 1^{m_2} 1^k B q_3 1^{m_2} 1 \rightarrow \cdots \rightarrow$$

$$B 1^{m_2} 1^k B 1^{m_2} q_3 B \rightarrow B 1^{m_2} 1^k B 1^{m_2} q_4 1 B \rightarrow B 1^{m_2} 1^k B 1^{m_2} q_4 B B \rightarrow \cdots$$

$$\cdots \rightarrow B 1^{m_2} 1^k B q_6 1^{m_2} B B \rightarrow B 1^{m_2} 1^k q_6 B 1^{m_2} B B \rightarrow \cdots$$

$$\rightarrow q_8 B 1^{m_2} 1^k B 1^{m_2} B B \rightarrow B q_1 1^m \; 1^k B 1^{m_2} B B = \alpha_s.$$

Exceptînd un B inițial și B final, α_s este o descriere instantanee de același tip cu α_1, dar cu o pereche de unități mai puțin. Procesul este acum repetat. Eventual,

$$\alpha_1 \rightarrow \cdots \alpha_s \rightarrow \cdots \rightarrow \ldots \rightarrow B^{m_2} q_1 1 1^k B 1 B^{m_2+1} \rightarrow \cdots \rightarrow$$

$$\rightarrow B^{m_2+1} 1^k q_2 B 1 B^{m_2+1} \rightarrow \cdots \rightarrow B^{m_2+1} 1^k B q_4 \, 1 B^{m_2+1} \rightarrow$$

$$B^{m_2+1} 1^k B q_4 \, B B^{m_2+1} \rightarrow B^{m_2+1} 1^k q_5 B B B^{m_2+1} = \alpha_p,$$

unde α_p este o descriere instantanee terminală. Însă $<\alpha_p> = k = m_1 \dot{-} m_2$, deci pentru $m_1 \geqslant m_2$,

$$\psi_Z^{(2)}(m_1, m_2) = m_1 \dot{-} m_2.$$

Însă aceasta nu este suficient pentru a decide că mașina Z considerată calculează funcția f. Mai trebuie arătat că pentru $m_1 < m_2$ funcția $\psi_Z^{(2)}$ nu este definită în punctul (m_1, m_2). În acest scop, să punem $k = m_2 - m_1$. Avem:

$$\alpha_1 = q_1 1^{m_1+1} B 1^{m_2+1} = q_1 11^{m_1} B 1^k 1^{m_1} 1 \to \cdots \to$$

$$Bq_1 1^m B 1^k 1^{m_1} BB \to \cdots \to \cdots \to B^{m_1} q_1 1 B 1^k 1 B^{m_1+1} \to$$

$$\to B^{m_1} q_1 B B 1^k 1 B^{m_1+1} \to B^{m_1} B q_2 B 1^k 1 B^{m_1+1} \to$$

$$\to B^{m_1} B B q_3 1^k 1 B^{m_1+1} \to \cdots \to B^{m_1} B B 1^k q_4 1 B^{m_1+1} \to$$

$$\to B^{m_1} B B 1^k q_4 B B^{m_1+1} \to \cdots \to B^{m_1} B q_6 B 1^k B B^{m_1+1} \to$$

$$\to B^{m_1} q_7 B B 1^k B B^{m_1+1} \to B^{m_1} B q_9 B 1^k B B^{m_1+1} = \alpha_7 \to$$

$$\to B^{m_1} B B q_9 1^k B B^{m_1+1} \to B^{m_1} B q_9 B 1^k B B^{m_1+1} \to \cdots.$$

Ultimele două descrieri instantanee sînt transformate, succesiv, una în cealaltă în mod indefinit, deci activitatea mașinii se perpetuează, fără să se ajungă la o descriere instantanee terminală. Rezultată că pentru $m_1 < m_2$, $\psi_Z^{(2)}(m_1, m_2)$ nu este definită. Deci

$$\psi_Z^{(2)}(x, y) = f(x, y) = x - y$$

și funcția f este parțial calculabilă în sensul lui Turing.

Am văzut că funcții dintre cele mai simple nu pot fi calculate decît cu ajutorul unor mașini Turing destul de complicate. Pentru a evita construirea directă a unor astfel de mașini, se utilizează diverse tehnici de compunere a unor mașini elementare standardizate. O mașină Turing este, prin definiție, *o mașină standardizată*, dacă este prevăzută cu *o stare finală* q_θ cu următoarele două proprietăți: 1) mașina Z se găsește la sfîrșitul oricărui calcul pe care-l efectuează în starea q_θ ; 2) nici o situație a lui Z nu comportă starea q_θ. Este comod ca θ să fie cel mai mare dintre indicii care servesc la numerotarea stărilor mașinii Z.

Procesul de standardizare a mașinilor Turing permite ca rezultatele (ieșirile) unei mașini să devină intrările (datele inițiale) ale altei mașini. Aceasta face posibilă o anumită operație de compunere a mașinilor Turing stan-

dardizate, operație care la rîndul ei permite stabilirea unor teoreme importante relative la operații cu funcții calculabile. Două operații cu astfel de funcții sînt cu deosebire importante:

a. *Operația de compunere.* Fie $g_1(y_1, \ldots, y_n), \ldots, g_m(y_1, \ldots, y_n)$ funcții cu valori numere naturale, fiecare funcție g_i fiind definită pe o parte D_i a lui $\mathcal{N}^n$. Fie f o funcție cu valori naturale, definită pe o parte D a lui $\mathcal{N}^m$. Numim compunere a funcțiilor $g_1, \ldots g_m$ și f funcția h dată de expresia

$$h(y_1, \ldots, y_n) = f[g_1(y_1, \ldots, y_n), \ldots, g_m(y_1, \ldots, y_n)],$$

definită pe o parte a mulțimii

$$\bigcap_{i=1}^{m} D_i.$$

Folosind mașini Turing standardizate, se poate arăta că dacă $f, g_1 \ldots, g_m$ sînt parțial calculabile (respectiv calculabile), atunci h este de asemenea parțial calculabilă (respectiv calculabilă) în sensul lui Turing. Cu alte cuvinte, clasa funcțiilor parțial calculabile este închisă în raport cu operația de compunere.

b. *Operația de minimizare.* Fie funcția $f(y, x_1, \ldots, x_n)$, definită în fiecare punct din $\mathcal{N}^{n+1}$ și avînd valori naturale. Acestei funcții îi asociem funcția $h(x_1, \ldots, x_n)$, a cărei valoare (dacă există) este, prin definiție, cea mai mică valoare (dacă există) a lui y, pentru care $f(y, x_1, \ldots, x_n) = 0$. Rezultă că funcția h este definită, în general, doar pe o parte a produsului cartezian $\mathcal{N}^n$. În cazul în care funcția $h(x_1, \ldots, x_n)$ asociată lui f este definită în fiecare punct din $\mathcal{N}^n$, spunem că $f(y, x_1, \ldots, x_n)$ este *regulată* (pentru y). Cu ajutorul mașinilor Turing standardizate, folosind o mașină care calculează succesiv pe $f(y, x_1, \ldots, x_n)$ pentru $y = 0$, $y = 1$, $y = 2, \ldots$, pînă cînd se obține $f(y, x_1, \ldots x_n) = 0$, se ajunge la următorul rezultat: dacă $f(y, x_1, \ldots, x_n)$ este parțial calculabilă, atunci funcția h, asociată lui f prin minimizare, este parțial calculabilă. Dacă, în plus, f este regulată, funcția h este calculabilă în sensul lui Turing.

Acum putem defini o clasă de funcţii deosebit de importantă în logica matematică, aceea a funcţiilor parţial recursive. Există diferite moduri de introducere a acestor funcţii (S.C. Kleene, Rosza Peter etc.), dar cel mai apropiat de teoria maşinilor Turing şi de noţiunile studiate mai sus este datorit lui J. Robinson. Vom adopta aici prezentarea acestui din urmă autor.

O funcţie definită pe o parte a lui $\mathcal{N}^n$ cu valori naturale este numită *funcţie parţial recursivă*, dacă ea se poate obţine prin aplicarea de un număr finit de ori a operaţiilor de compunere şi de minimizare, pornind de la următoarele funcţii de bază:

(1) funcţia succesor $S(x) = x + 1$;
(2) funcţia de proiecţie $U_i^n(x_1,\ldots,x_n) = x_i (1 \leqslant i \leqslant n)$;
(3) funcţia de adunare $x + y$;
(4) funcţia de scădere cu prelungire trivială

$$x - y = \begin{cases} x - y, \text{ dacă } x \geqslant y; \\ 0, \text{ dacă } x < y; \end{cases}$$

(5) funcţia produs xy.

Despre funcţiile (1) şi (3) s-a arătat deja că sînt calculabile în sensul lui Turing. Din faptul că, aşa cum s-a arătat, funcţia de scădere $x - y$ este calculabilă, rezultă fără dificultate că şi funcţia de scădere cu prelungire trivială este calculabilă Turing. Se poate arăta că şi funcţiile (2) şi (5) sînt calculabile Turing. Deoarece operaţiile de compunere şi minimizare, aplicate unor funcţii calculabile Turing, conduc tot la funcţii calculabile Turing, rezultă următoarea teoremă deosebit de importantă:

Orice funcţie parţial recursivă este parţial calculabilă în sensul lui Turing.

Se poate arăta că şi reciproca acestei teoreme este adevărată; cu alte cuvinte, clasa funcţiilor parţial recursive coincide cu clasa funcţiilor calculabile în sensul lui Turing.

O funcţie parţial recursivă este, prin definiţie, o *funcţie recursivă*, dacă operaţiile care o definesc se aplică exclusiv unor funcţii regulate. Se poate arăta că există identitate între clasa funcţiilor recursive şi clasa funcţiilor calculabile în sensul lui Turing.

Ţinînd seamă că noţiunea de algoritm normal este echivalentă cu aceea de funcţie parţial calculabilă în sensul lui Turing, rezultă că dispunem de trei variante echivalente pentru conceputul de calculabilitate. Acestea nu sînt singurele, dar sînt, în orice caz, cele mai importante. Pe una sau alta dintre aceste variante se poate funda analiza matematică a funcţiilor constructive, şi unele idei în această privinţă au fost deja schiţate într-un paragraf precedent.

II

Toate tipurile de trecere la limită admit o schemă comună

INTRODUCERE

Procesul de trecere la limită apare în matematică sub formele cele mai variate. De la forma rudimentară pe care o prezintă noțiunea de „șir convergent", pînă la forma mult mai complexă pe care o prezintă ideea de integrală există un mare număr de nuanțe, de diverse grade de complexitate ale unui același proces: trecerea la limită. Pe de altă parte dat fiind că de obicei procesele infinite care apar în matematică oglindesc un conținut intuitiv care ne este familiar, se întîmplă de multe ori să ne lăsăm furați de acest aspect intuitiv și să utilizăm operații de trecere la limită într-o formă care nu are drept de cetate în matematică; într-o formă în care nu se înțelege bine despre ce este vorba, deci care lasă de dorit din punctul de vedere al rigorii.

În acest capitol ne propunem să arătăm că toate tipurile de trecere la limită pe care le întîlnim în matematică, oricît de diferite ar putea ele să pară, pot fi prezentate dintr-un punct de vedere unic, deci ca aspecte ale unui același proces. Faptul acesta nu a fost cunoscut de către vechii analiști, el este o cucerire a secolului nostru. Acel proces unic care conține drept cazuri particulare toate variantele posibile de trecere la limită, poate fi ales în două feluri. Noi vom expune aici o metodă care se sprijină pe o noțiune introdusă în matematică în urmă cu vreo treizeci de ani, de către *Henri Cartan*, noțiunea de filtru. Celălalt mod de tratare a proceselor de trecere la limită se găsește expus în tratatul de *Analiză matematică* al acad. Miron Nicolescu. A se vedea, în special, volumul

al doilea al acestui tratat. (Editura Tehnică, București, 1958).

Iată, mai întîi, o listă (incompletă) de procese la limită:

limita unui șir;
marginea superioară sau inferioară a unei mulțimi liniare;
limita superioară sau inferioară a unei funcții definite pe o mulțime;
marginea superioară sau inferioară a unei funcții într-un punct;
limita superioară sau inferioară a unei funcții într-un punct;
limita unei funcții într-un punct;
limita superioară sau inferioară a unui șir;
variația totală a unei funcții pe un interval;
oscilația totală a unei funcții pe un interval;
integralele Darboux ale unei funcții pe un interval.

Noțiunea de filtru, pe care o vom expune imediat, are un caracter foarte abstract. În aparență, n-are nici o legătură cu procesele la limită înșirate mai sus. Expunerea definiției și proprietăților acestei noțiuni este destul de aridă. Această ariditate însă (care nu se va prelungi mult) va fi pe deplin compensată de satisfacția pe care o vom avea cînd vom constata că noțiunea nou introdusă lămurește esența tuturor proceselor infinite pe care le-am întîlnit pînă acum în analiză. Să trecem la definirea ei.

CE ESTE UN FILTRU?

Fie E o mulțime nevidă de elemente de natură nespecificată. Fie $\mathcal{F}$ o mulțime de părți ale lui E, avînd următoarele proprietăți:

F_1) dacă $M \in \mathcal{F}$, iar $M \subset N \subset E$, atunci $N \in \mathcal{F}$;

F_2) dacă $M \in \mathcal{F}$, $N \in \mathcal{F}$, atunci $M \cap N \in \mathcal{F}$;

F_3) partea vidă a lui E nu aparține lui $\mathcal{F}$.

$\mathcal{F}$ se numește un filtru asupra lui E.

Iată un exemplu de filtru: mulțimea tuturor părților dreptei care conține un același punct x este un filtru asupra lui R^1. În adevăr, dacă o parte a dreptei conține pe x,

atunci orice parte mai cuprinzătoare conține de asemenea pe x. Intersecția a două părți care conțin pe x este o parte care conține pe x. În sfîrșit, partea vidă a dreptei nu conține pe x.

Mulțimile acestui filtru sînt, într-un limbaj familiar, vecinătățile punctului x. Este suficient să ținem seama de importanța vecinătăților în definiția noțiunii de limită, pentru ca să începem să bănuim legătura dintre filtre și limite.

Vom da acum cîteva noțiuni ajutătoare și unele proprietăți ale filtrelor.

Un filtru $\mathcal{F}'$ este mai fin decît un filtru $\mathcal{F}$ asupra aceleiași mulțimi E, dacă $\mathcal{F} \subset \mathcal{F}'$.

O mulțime $\mathcal{B}$ de părți ale lui E este o bază de filtru, dacă satisface următoarele condiții:

B_1) dacă $B' \in \mathcal{B}$, $B'' \in \mathcal{B}$, atunci există $B \in \mathcal{B}$, așa încît $B \subset B' \cap B''$;

B_2) $\mathcal{B}$ nu este vidă;

B_3) $\mathcal{B}$ nu conține partea vidă a lui E.

Pentru a explica această definiție, să demonstrăm.

Teorema 1. *Mulțimea $\mathcal{F}$, compusă din toate părțile lui E care conțin, fiecare, o mulțime a lui $\mathcal{B}$, este un filtru asupra lui E.*

Într-adevăr, orice parte A a lui E care conține o mulțime F a lui $\mathcal{F}$ conține, prin definiție, o mulțime $B \subset F$ a lui $\mathcal{B}$, deci aparține lui $\mathcal{F}$ și, prin urmare, $\mathcal{F}$ satisface F_1.

Intersecția $F \cap F'$ a două mulțimi ale lui $\mathcal{F}$ conține intersecția $B \cap B'$ a mulțimilor $B \subset F$, $B' \subset F'$ ale lui $\mathcal{B}$; însă conform condiției B_1, $B \cap B'$ conține o mulțime $B'' \in \mathcal{B}$, deci $B'' \subset F \cap F'$ și, prin urmare, $F \cap F' \in \mathcal{F}$. Este satisfăcută deci F_2.

În sfîrșit, $\mathcal{F}$ nu conține partea vidă a lui E, deoarece $\mathcal{B}$ nu conține această mulțime, deci $\mathcal{F}$ satisface F_3.

$\mathcal{F}$ este deci un filtru asupra lui E; se numește filtru generat de baza $\mathcal{B}$.

Două baze de filtru $\mathcal{B}$ și $\mathcal{B}'$ sînt echivalente dacă generează același filtru.

Teorema 2. *Pentru ca un filtru $\mathcal{F}'$ de bază $\mathcal{B}'$ să fie mai fin decît un filtru $\mathcal{F}$ de bază $\mathcal{B}$, este necesar și suficient ca fiecare mulțime B a lui $\mathcal{B}$ să conțină o mulțime $B' \in \mathcal{B}'$.*

Demonstrație. Într-adevăr, dacă $\mathcal{F}' \supset \mathcal{F}$, atunci fiecare mulțime F a lui $\mathcal{F}$ aparține lui $\mathcal{F}'$. În particular fiecare mulțime B a lui $\mathcal{B}$ aparține lui $\mathcal{F}'$, deoarece $\mathcal{B} \subset \mathcal{F}$, deci $\mathcal{B}$ conține o mulțime $B' \in \mathcal{B}'$.

Reciproc, dacă fiecare mulțime B a lui $\mathcal{B}$ conține o mulțime B' a lui $\mathcal{B}'$, atunci $\mathcal{B} \subset \mathcal{F}'$, deci, în virtutea lui F_2, $\mathcal{F} \subset \mathcal{F}'$.

Fie în R^n mulțimea $S_r(x)$ a punctelor y situate la o distanță inferioară lui r de punctul x ($r > 0$).

Mulțimea $\mathcal{S}(x)$ a tuturor sferelor cu centrul în x constituie o bază de filtru. În adevăr, intersecția a două sfere $S_r(x)$ și $S_{r'}(x)$ de centru x și de raze r și r' este sfera cu raza cea mai mică, deci $\mathcal{S}(x)$ satisface B_1.

$\mathcal{S}(x)$ nu este vidă și nu conține partea vidă din R^n (deoarece $r > 0$), deci $\mathcal{S}(x)$ satisface B_2 și B_3.

Filtrul $\mathcal{V}(x)$ de bază $\mathcal{S}(x)$, adică mulțimea tuturor părților lui R^n care conțin fiecare o sferă $S_r(x)$, poartă numele de filtrul vecinătăților punctului x; fiecare dintre mulțimile $V(x)$ ale lui $\mathcal{V}(x)$ se numește vecinătate a lui x.

Orice bază a filtrului vecinătăților poartă numele de sistem fundamental de vecinătăți, iar mulțimile bazei numele de vecinătăți fundamentale: în particular, sferele $S_r(x)$ sînt vecinătăți fundamentale ale lui x. În plan, așa-numitele sfere se confundă cu discurile centrate în x. Pe dreaptă, sferele se confundă cu intervalele centrate în x.

Un filtru $\mathcal{F}$ asupra spațiului R^n se spune că are drept limită (converge către) un punct x, dacă $\mathcal{F}$ este mai fin decît filtrul $\mathcal{V}(x)$ al vecinătăților lui x.

O bază de filtru $\mathcal{B}$ are ca limită (converge către) un punct x, dacă filtrul $\mathcal{F}$ de bază $\mathcal{B}$ converge către x.

Teorema 3. *Condiția necesară și suficientă ca o bază de filtru $\mathcal{B}$ să conveargă către x este ca fiecare vecinătate fundamentală a unui sistem fundamental oarecare de vecinătăți ale lui x să conțină o mulțime a lui $\mathcal{B}$.*

Demonstrație. Dacă $\mathcal{B}$ satisface aceste condiții, atunci, în virtutea teoremei 2, filtrul $\mathcal{F}$ de bază $\mathcal{B}$ este mai fin decît filtrul $\mathcal{V}(x)$ al vecinătăților lui x. Reciproc, dacă $\mathcal{F} \supset \mathcal{V}(x)$, atunci, în virtutea aceleiași teoreme 2, fiecare vecinătate fundamentală a lui x conține o mulțime a lui $\mathcal{B}$.

Teorema 4. *În spațiul R^n orice filtru convergent are o limită unică.*

În adevăr, în caz contrar, un filtru $\mathcal{F}$ ar tinde cel puțin către două puncte distincte x și y, deci $\mathcal{F} \supset \mathcal{V}(x), \mathcal{F} \supset \mathcal{V}(y)$.

Fie sferele $S_r(x)$ și $S_r(y)$. Avem $S_r(x) \in \mathcal{V}(x)$, $S_r(y) \in \mathcal{V}(y)$

Fie d distanța dintre x și y și fie $2r < d$. Rezultă $S_r(x) \cap S_r(y) = 0$. Ținînd seama că aceste două sfere aparțin filtrului $\mathcal{F}$, am ajuns la o contradicție care demonstrează teorema.

Vom arăta acum că toate procesele infinite pe care le-am întîlnit și pe care le-am amintit la începutul acestui capitol pot fi reduse la o singură idee: aceea de convergență a unui anumit filtru.

MARGINEA SUPERIOARĂ A UNEI MULȚIMI LINIARE DE PUNCTE

Fie o mulțime $\mathcal{G}$ liniară, nevidă. Pentru fiecare $x \in \mathcal{G}$ considerăm partea $A_x \subset \mathcal{G}$ definită astfel: $x' \in A_x$ dacă și numai dacă $x' \in E$ și $x' \geqslant x$. Mulțimea părților $\{A_x\}$ cînd x parcurge pe $\mathcal{G}$ formează o bază de filtru $\mathcal{B}$; în adevăr, 1) $A_x \cap A_y$ conține orice A_z cu $z \in \mathcal{G}$; $z \geqslant \max(x, y)$; 2) $\mathcal{B}$ nu este vidă, deoarece $\mathcal{G}$ nu este vidă; există $x \in \mathcal{G}$, deci există $A_x \in \mathcal{B}$; 3) $\mathcal{B}$ nu conține partea vidă a lui $\mathcal{G}$, deoarece orice A_x conține cel puțin un element: pe x.

Fie $\mathcal{F}$ filtrul asupra lui R^1, care are ca bază pe $\mathcal{B}$. Avem:

Teoremă. *Necesar și suficient ca M să fie marginea superioară a lui $\mathcal{G}$ este ca*

$$\lim \mathcal{F} = M.$$

Demonstrație. Vom presupune că $\mathcal{G}$ este mărginită superior.

Necesitatea. Fie M marginea superioară a lui $\mathcal{G}$. Fie o vecinătate fundamentală a lui M: $(M - \varepsilon, M + \varepsilon)$ $(\varepsilon > 0)$. Conform definiției lui M, există $x \in \mathcal{G}$, astfel încît $M - \varepsilon < x$ și avem, în orice caz, $x \leqslant M$; deci $A_x \subset (M - \varepsilon, M + \varepsilon)$.

De aici rezultă imediat că $\mathcal{F}$ este mai fin ca $\mathcal{V}(M)$, conform teoremei 3.

Suficiența. Fie $\lim \mathcal{F} = M$. Fie $\varepsilon > 0$ și fie intervalul (vecinătatea fundamentală) $M - \varepsilon$, $M + \varepsilon$). Există $x \in \mathcal{G}$, astfel încît $A_x \subset (M - \varepsilon, \ M + \varepsilon)$ (conform teoremei 3). Însă A_x conține toate punctele din $\mathcal{G}$ situate la dreapta lui x, deci $x \in \mathcal{G} \rightarrow x < M + \varepsilon$. Deoarece ε a fost ales arbitrar, rezultă că $x \in \mathcal{G} \rightarrow \ x \leqslant M$.

Pe de altă parte, din $A_x \subset (M - \varepsilon, \ M + \varepsilon)$ rezultă că pentru orice $\varepsilon > 0$ există $x \in E$, astfel încît $M - \varepsilon < x$. M este, deci, marginea superioară a lui $\mathcal{G}$.

Observația 1. Teorema se extinde și la cazul în care $M = \infty$.

Observația 2. O teoremă analogă se poate da pentru marginea inferioară a mulțimii $\mathcal{G}$.

LIMITA SUPERIOARĂ A UNEI MULȚIMI LINIARE INFINITE

Fie $\mathcal{G}^*$ submulțimea acelor puncte din $\mathcal{G}$ la dreapta cărora există o infinitate de puncte din $\mathcal{G}$. Presupunem, deocamdată, că $\mathcal{G}^*$ nu e vidă. Fie, pentru $x \in G^*$, A_x^* formată din toate punctele $x' \in \mathcal{G}$ cu $x' \geqslant x$.

Mulțimea $\mathcal{B}^*$ a părților A_x^* ale lui $\mathcal{G}$ formează o bază de filtru. Într-adevăr, $A_x^* \cap A_y^*$ conține pe A_z^* cu $z = \max(x, y)$; $\mathcal{B}^*$ nu este vidă, deoarece $\mathcal{G}^*$ nu este vidă; $\mathcal{B}^*$ nu conține partea vidă a lui $\mathcal{G}$, deoarece pentru orice $x \in \mathcal{G}^*$, A_x^* nu este vidă (este chiar infinită).

Fie F^* filtrul de bază $\mathcal{B}^*$ Printr-un raționament analog cu cel precedent, se constată că

$$\lim \mathcal{F}^* = L,$$

unde L este limita superioară a lui $\mathcal{G}$.

Dacă $\mathcal{G}^*$ este vidă, atunci considerăm submulțimea ${}^*\mathcal{G}$ a acelor puncte din $\mathcal{G}$ la stînga cărora există o infinitate de puncte din $\mathcal{G}$ (${}^*\mathcal{G}$ va coincide, în acest caz, cu $\mathcal{G}$) și obținem un filtru ${}^*\mathcal{F}$ cu

$$\lim {}^*\mathcal{F} = L = l,$$

unde l este limita inferioară a lui $\mathcal{G}$.

MARGINEA SUPERIOARĂ A UNEI FUNCȚII DEFINITE PE O MULȚIME DIN R^n

Pentru fiecare $x \in \mathcal{G}$ considerăm mulțimea A_x a tuturor punctelor $x' \in \mathcal{G}$, pentru care $f(x') \geqslant f(x)$.

Mulțimea $\mathcal{B}$ a părților A_x ale lui $\mathcal{G}$ este o bază de filtru asupra lui $\mathcal{G}$. Într-adevăr, $A_x \cap A_y \supset A_z$, unde $z \in \mathcal{G}$ și $f(z) \geqslant \max(f(x), f(y))$. $\mathcal{B}$ nu este vidă, deoarece $\mathcal{G}$ nu este vidă; $\mathcal{B}$ nu conține partea vidă a lui $\mathcal{G}$, deoarece, pentru orice $x \in \mathcal{G}$, A_x conține cel puțin elementul x.

Fie $\mathcal{F}$ filtrul asupra lui $\mathcal{G}$, generat de $\mathcal{B}$.

Teoremă. *Necesar și suficient ca M să fie marginea superioară a lui f pe $\mathcal{G}$ este ca*

$$M = \lim_{\mathcal{F}} f.$$

Să presupunem că M este finit.

Necesitatea. Fie M marginea superioară a lui f pe $\mathcal{G}$. Avem deci

$$x \in \mathcal{G} \rightarrow f(x) \leqslant M$$

și pentru $\varepsilon > 0$ există $x \in \mathcal{G}$, așa încît $M - \varepsilon < f(x)$.

Cu alte cuvinte, pentru orice vecinătate fundamentală $(M - \varepsilon, M + \varepsilon)$ a lui M există $x \in \mathcal{G}$, astfel încît

$$f\{A_x\} \subset (M - \varepsilon, M + \varepsilon).$$

Însă $f\{A_x\}$ aparține bazei de filtru $f(\mathcal{F})$ asupra lui R^1. Deci conform teoremei 3, $f(\mathcal{F})$ converge către M.

Suficiența. Fie $N = \lim_{\mathcal{F}} f$. Dacă

$$\sup f(x) = M \neq N,$$

atunci, conform demonstrației precedente, $\lim_{\mathcal{F}} f = N$. Însă aceasta ar contrazice teorema 4 asupra unicității limitei unui filtru asupra lui R^n.

Observația 1. Teorema rămîne în vigoare, cu mici modificări în demonstrație, în cazul $M = +\infty$.

Observația 2. O teoremă analogă se obține pentru marginea inferioară a lui f pe $\mathcal{G}$; A_x se înlocuiește cu A'_x conținînd toate punctele $x' \in \mathcal{G}$ pentru care $f(x') \leqslant f(x)$.

Fie $f(x)$ local mărginită superior[1] ($x \in R^1$). Fie $x_0 \in R^1$. Prin definiție:

$$M(x_0) = \sup\ (f;\ x_0) = \inf\ (\sup\ \{f;\ I_x\},$$

unde $I_x = [2x_0 - x,\ x]$.

Introducem funcția μ definită pe R^1 în modul următor:

$$\mu(x) = \mu(2x_0 - x) = \sup\ (f;\ I_x).$$

Din definiția marginii superioare în x_0, rezultă că

$$M(x_0) = \inf_{x \in R^1}\ \mu(x).$$

Mulțimile A'_x (vezi observația 2 de mai sus) sînt aici chiar intervalele închise I_x. Baza de vecinătăți formată de intervalele închise I_x generează un filtru $\mathcal{F}$ asupra lui R^1. (De observat că acest filtru este mai fin ca $\mathcal{V}(x_0)$; dacă, însă, am fi definit funcția μ numai pentru $x \neq x_0$, am fi obținut $\mathcal{F} = \mathcal{V}(x_0)$.) Avem

$$\lim_{\mathcal{F}} \mu = M(x_0)$$

și, conform observației din paranteză, avem chiar

$$\lim_{\mathcal{V}x_0} \mu = M(x_0).$$

Observația 1. Luînd

$$\mu^*(x) = \mu^*(2x_0 - x) = \inf\ (f;\ I_x),$$

obținem

$$\lim_{\mathcal{F}} \mu^* = \lim_{\mathcal{V}x_0} \mu^* = m(x_0)\ \text{(marg. inf. în } x_0).$$

Observația 2. Să presupunem că există N, astfel încît $\lim_{\mathcal{F}} \mu = N$. ($\mathcal{F}$ fiind filtrul definit mai sus). Conform teoremei de unicitate a limitei unui filtru în R^n, rezultă $N = M(x_0)$.

---

1 O funcție f este local mărginită superior, dacă pentru fiecare punct există o vecinătate în care f este mărginită superior. Este vizibil că f poate fi local mărginită superior, fără a fi mărginită superior, cum se întîmplă cu $f(x) = \frac{1}{x}$ în (0, 1).

LIMITA SUPERIOARĂ (RESPECTIV INFERIOARĂ) A UNEI FUNCȚII ÎNTR-UN PUNCT

Fie f definită și local mărginită pe R. (Condiția mărginirii locale este cerută numai pentru a evita introducerea lui $-\infty$ și $+\infty$).

Fie o funcție v, definită pentru orice $x \neq x_0$ în modul următor:

$$v(x) = v(2x_0 - x) = \sup\,(f;\ I_x - \{x_0\}),\quad (I_x = [2x_0 - x,\ x]).$$

Deoarece, prin definiție,

$$\overline{\lim_{x \to x_0}}\ f(x) = \inf_{x \neq x_0}\,(\sup\,[f;\ I_x - \{x_0\}]),$$

rezultă, ca și mai sus, că

$$\lim\nolimits_{\mathfrak{F}x_0} v = \overline{\lim_{x \to x_0}}\ f(x).$$

Observația 1. O teoremă analogă se obține pentru $\lim\limits_{x \to x_0} f(x)$, înlocuindu-se funcția v cu funcția v^*, definită astfel:

$$v^*(x) = v^*(2x_0 - x) = \inf\,(f;\ I_x - \{x_0\}).$$

Observația 2. Unicitatea limitelor

$$\lim\nolimits_{\mathfrak{F}(x_0)}\,(x_0)\ v,\quad \lim\nolimits_{\mathfrak{F}(x_0)}\ v^*$$

rezultă din teorema de unicitate a limitei unui filtru asupra lui R^n.

LIMITA UNEI FUNCȚII ÎNTR-UN PUNCT

Fie f definită în R^1, cu valori în R^1.

Teoremă. *O condiție necesară și suficientă ca să existe* $\lim\limits_{x \to x_0} f(x)$ *este să se poată defini o funcție* φ, *continuă în* x_0 *și egală cu* $f(x)$ *pentru* $x \neq x_0$, *astfel încît să existe* $\lim\nolimits_{\mathfrak{F}(x_0)}\varphi$. Avem:

$$\lim_{x \to x_0} f(x) = \lim\nolimits_{\mathfrak{F}(x_0)}\varphi.$$

Demonstrație. Necesitatea. Presupunem că $\lim_{x \to x_0} f(x) = A$. Fiind dat $\varepsilon > 0$, există $\eta_\varepsilon > 0$, astfel încît $x \in (x_0 - \eta_\varepsilon, x_0 + \eta_\varepsilon)$ să rezulte

$$\varphi(x) \in (A - \varepsilon, \quad A + \varepsilon)$$

sau

$$\varphi\{(x_0 - \eta_\varepsilon, \quad x_0 + \eta_\varepsilon)\} \subset (A - \varepsilon, \quad A + \varepsilon),$$

deci orice vecinătate fundamentală a lui A conține o mulțime din baza de filtru $\varphi(\mathcal{V}(x_0))$, cu alte cuvinte, această bază de filtru converge către A.

Suficiența. Presupunem că $\varphi(\mathcal{B}) \to A$; deci, orice vecinătate fundamentală $(A - \varepsilon, A + \varepsilon)$ conține o mulțime din $\varphi(\mathcal{V}(x_0))$, deci o mulțime din $\varphi(\mathcal{S}(x_0))$; există deci un $\eta > 0$, așa încît

$$\varphi(x_0 - \eta, \quad x_0 + \eta) \subset (A - \varepsilon, \quad A + \varepsilon),$$

ceea ce este echivalentă cu

$$\lim_{x \to x_0} f(x) = A.$$

LIMITA UNUI ȘIR

Fie un șir cu termeni reali, adică o funcție definită pe $\mathcal{N}$ (mulțimea numerelor naturale), cu valori în R^1. Vom nota această funcție cu f, iar valoarea lui f în n cu $f(n)$.

Fie $B_n = (n, n + 1, n + 2, ..)$. B_n se numește o secțiune superioară în $\mathcal{N}$. Fie $\mathcal{B}$ mulțimea tuturor secțiunilor superioare ale lui $\mathcal{N}$. $\mathcal{B}$ constituie o bază de filtru asupra lui $\mathcal{N}$. În adevăr, $B_n \cap B_m \supset B_p$ cu $p \geqslant \max(m, n)$; $\mathcal{B}$ nu este vidă, căci $\mathcal{N}$ nu este vidă și fiecare $n \in \mathcal{N}$ determină un B_n; $\mathcal{B}$ nu conține partea vidă a lui $\mathcal{N}$, căci orice B este diferit de mulțimea vidă. Fie $\mathcal{F}$ filtrul generat de $\mathcal{B}$ asupra lui $\mathcal{N}$. Avem

T e o r e m ă. *Necesar și suficient ca*

$$\lim_{n \to \infty} f(n) = A$$

este ca

$$\lim_{\mathcal{F}} f = A.$$

Necesitatea. Presupunem că

$$\lim_{n\to\infty} f(n) = A,$$

deci că pentru $\varepsilon > 0$ există un număr natural N_ε, astfel încît

$$n > N_\varepsilon \Rightarrow f(n) \in (A - \varepsilon, \quad A + \varepsilon).$$

Urmează că o vecinătate fundamentală $(A - \varepsilon, A + \varepsilon)$ a lui A conține mulțimea

$$f(B_{N_\varepsilon}) = (f(N_\varepsilon), \quad f(N_\varepsilon + 1), \quad f(N_\varepsilon + 2), \ldots.),$$

care este un element din baza de filtru $f(\mathcal{B}) = \{f(B_n)\}$. Rezultă deci că baza de filtru $f(\mathcal{B})$ converge către A, deci că baza de filtru $f(\mathcal{F})$ converge către A și avem

$$\lim_{\mathcal{F}} f = A.$$

Suficiența. Fie $\lim_{\mathcal{F}} f = A$; urmează că baza de filtru $f(\mathcal{F})$ converge către A, deci că orice vecinătate fundamentală $(A - \varepsilon, A + \varepsilon)$ conține o mulțime din $f(\mathcal{F})$ și, cu atît mai mult, din $f(\mathcal{B})$; fie această mulțime $f(B_{N_\varepsilon})$. Din

$$f(B_{N_\varepsilon}) \subset (A - \varepsilon, \quad A + \varepsilon)$$

rezultă imediat

$$\lim_{n\to\infty} f(n) = A.$$

LIMITA SUPERIOARĂ (RESPECTIV INFERIOARĂ) A UNUI ȘIR

Limita superioară a unui șir a fost definită ca marginea superioară a mulțimii valorilor care se obțin ca limite ale diferitelor subșiruri ale șirului dat[1].

---

[1] S-ar putea întîmpla ca această margine superioară să fie $+\infty$. Nu este greu de văzut că tot ceea ce urmează rămîne valabil și în acest caz. La nevoie, șirul $\mu(1), \mu(2), \ldots, \mu(n), \ldots$ Va deveni $+\infty$, $+\infty+, \ldots +\infty, \ldots$

Însă un șir este o funcție f definită pe $\mathcal{N}$, cu valori în R^1, deci ar trebui ca definiția de mai sus să corespundă noțiunii de limită superioară a unei funcții într-un punct. Faptul că acest punct este $+\infty$ nu schimbă esențial situația. „Vecinătățile" fundamentale ale lui $+\infty$ sînt, în mulțimea $\mathcal{N}$, secțiunile superioare B_n. Nu este greu de arătat că

$$\overline{\lim_{x\to\infty}} f(n) = \inf_n (\sup \{f(n), f(n+1), \ldots\}).$$

Introducînd funcția auxiliară μ, definită pe $\mathcal{N}$ prin

$$\mu(n) = \sup\{f(n), f(n+1),..\} = \sup (f; B_n)$$

și ținînd seamă că șirul numeric

$$\mu(1), \quad \mu(2), \ldots, \mu(n),\ldots$$

este monoton descrescător, deci că are limită, rezultă:

$$\overline{\lim_{n\to\infty}} f(n) = \lim_{n\to\infty} \mu(n).$$

Dar, după cum am văzut,

$$\lim_{n\to\infty} \mu(n) = \lim_{\mathcal{F}} \mu,$$

unde $\mathcal{F}$ este filtrul asupra lui $\mathcal{N}$, generat de baza $\mathcal{B}$ a secțiunilor $\{B_n\}$. Urmează că

$$\overline{\lim_{n\to\infty}} f(n) = \lim_{\mathcal{F}} \mu.$$

În mod analog se poate arăta că

$$\underline{\lim_{n\to\infty}} f(n) = \lim_{\mathcal{F}} \nu$$

unde ν este o funcție reală definită pe $\mathcal{N}$ în modul următor:

$$\nu(n) = \inf (f; B_n).$$

VARIAȚIA TOTALĂ A UNEI FUNCȚII

Fie f o funcție definită pe $[a, b]$. Să definim pe mulțimea Δ a diviziunilor lui $[a, b]$ o funcție v în felul următor: dacă $\Delta \in \Delta$.

$$v(\Delta) = \sum_{i=0}^{n-1} |f(x_{i+1}) - f(x_i)|,$$

unde $\Delta = (a = x_0, \ldots, x_n = b)$.

Să considerăm, pentru fiecare $\Delta \in \Delta$, mulțimea $\mathcal{S}_\Delta$ a tuturor diviziunilor lui $[a, b]$ mai fine ca Δ. $\mathcal{S}_\Delta$ se numește o secțiune la dreapta în Δ.

Mulțimea $\mathcal{S}$ a secțiunilor la dreapta în Δ este o bază de filtru asupra lui Δ. În adevăr, $\mathcal{S}_{\Delta'} \cap \mathcal{S}_{\Delta''} \supset \mathcal{S}_\Delta$ dacă $\Delta \supset \Delta' \cup \Delta''$; $\mathcal{S}$ nu este vidă, deoarece fiecare diviziune a lui $[a, b]$ induce o secțiune la dreapta în Δ; $\mathcal{S}$ nu conține partea vidă a lui Δ, deoarece orice $\mathcal{S}_\Delta$ conține o infinitate de elemente din Δ.

Fie $\mathcal{F}$ filtrul asupra lui Δ, generat de $\mathcal{S}$. Avem

Teoremă. *Necesar și suficient ca f să fie cu variație mărginită pe $[a, b]$ este să existe și să fie finită limita*

$$\lim_{\mathcal{F}} v.$$

Valoarea acestei limite este variația totală a lui f, de la a la b.

Necesitatea. Fie $V = \overset{b}{\underset{a}{V}}(f)$. Fie $\varepsilon > 0$. Există o diviziune Δ_ε, astfel încît $V - \varepsilon < v(\Delta_\varepsilon)$. Fie acum $\Delta \supset \Delta_\varepsilon$. Avem, cu atît mai mult, $V - \varepsilon < v(\Delta)$. Deoarece avem și $v(\Delta) \leqslant V$, rezultă că orice vecinătate fundamentală a lui V conține imaginea, prin funcția v, a unei secțiuni $\mathcal{S}_\Delta$. Deci baza de filtru $v(\mathcal{F})$ converge către V, adică

$$\lim_{\mathcal{F}} v = V.$$

Suficiența. Din $\lim_{\mathcal{F}} v = V$ rezultă că $(V - \varepsilon, V + \varepsilon)$ conține o mulțime din baza de filtru $v(\mathcal{F})$, deci și o mulțime din $v(\mathcal{S})$, adică imaginea prin v a unei anumite secțiuni $\mathcal{S}_{\Delta_\varepsilon}$. Pentru orice $\Delta \supset \Delta_\varepsilon$ avem

$$V - \varepsilon < v(\Delta) < V + \varepsilon.$$

De aici rezultă că

$$V \leqslant \overset{b}{\underset{a}{V}}(f).$$

Fie acum o diviziune arbitrară Δ și fie $\Delta_1 = \Delta_\varepsilon \cup \Delta$. Din $\Delta_1 \supset \Delta_\varepsilon$ rezultă

$$v(\Delta_1) < V + \varepsilon,$$

deci, cu atît mai mult,

$$v(\Delta) < V + \varepsilon.$$

(Se folosește proprietatea că $v(\Delta)$ nu descrește cînd se trece la o diviziune mai fină).

Deoarece ε este un număr pozitiv arbitrar, iar Δ este un element arbitrar din $\mathbf{\Delta}$, rezultă că pentru orice Δ

$$v(\Delta) \leqslant V,$$

deci că f este cu variație mărginită pe $[a, b]$. Avem:

$$\overset{b}{\underset{a}{V}}(f) \leqslant V.$$

Deoarece am obținut și o inegalitate de sens contrar avem

$$\overset{b}{\underset{a}{V}}(f) = V.$$

Observația 1. Funcția v poate fi înlocuită cu funcția ω, definită în felul următor: dacă $\Delta \in \mathbf{\Delta}$.

$$\omega(\Delta) = \sum_{i=0}^{n-1} \omega(f;\ [x_i,\ x_{i+1}]).$$

Observația 2. Să considerăm, pentru fiecare $\Delta \in \mathbf{\Delta}$ mulțimea $\mathcal{S}^*_\Delta$ a diviziunilor de normă $\leqslant ||\Delta||$ (prin $||\Delta||$ notăm norma diviziunii Δ).

Mulțimile $\mathcal{S}^*_\Delta$ formează o bază de filtru $\mathcal{S}^*$. În adevăr, $\mathcal{S}^*_{\Delta'} \cap \mathcal{S}^*_{\Delta''} \supset \mathcal{S}^*_\Delta$, unde $||\Delta|| < \min(||\Delta'||, ||\Delta''||)$. $\mathcal{S}^*$ nu este vidă, deoarece $\mathbf{\Delta}$ nu este vidă și fiecare $\Delta \in \mathbf{\Delta}$ generează un $\mathcal{S}^*_\Delta$; $\mathcal{S}^*$ nu conține partea vidă a lui $\mathbf{\Delta}$, deoarece orice $\mathcal{S}^*_\Delta$ conține o infinitate de

elemente din Δ. Fie $\mathcal{F}^*$ filtrul generat de $\mathcal{S}^*$ ($\mathcal{S}^*_\Delta$ este tot o secțiune la dreapta în Δ, în care diviziunile sînt ordonate după normă, spre deosebire de $\mathcal{S}_\Delta$, unde diviziunile sînt ordonate după finețe).

O funcție f poate fi cu variație mărginită pe $[a, b]$ și totuși să nu existe

$$\lim_{\mathcal{F}^*} v.$$

Se poate arăta însă că dacă f este continuă și cu variație mărginită, atunci

$$\overset{b}{\underset{a}{V}}(f) = \lim_{\mathcal{F}} v = \lim_{\mathcal{F}^*} v.$$

Într-adevăr, aceasta revine la proprietatea binecunoscută că fiind dată o funcție f continuă și cu variație mărginită, pentru orice șir Δ_n de diviziuni de normă tinzînd la zero avem

$$\lim_{n\to\infty} v(\Delta_n) = \overset{b}{\underset{a}{V}}(f).$$

Există și funcții discontinue, cu variație mărginită, pentru care

$$\overset{b}{\underset{a}{V}}(f) = \lim_{\mathcal{F}} v = \lim_{\mathcal{F}^*} v;$$

sînt toate funcțiile cu variație mărginită pentru care avem, pentru orice x,

$$\min(f(x-0), f(x+0)) \leqslant f(x) < \max(f(x-0), f(x+0)).$$

Să considerăm funcția ω, definită pe Δ, așa cum s-a arătat mai sus. Vom arăta că pentru orice funcție f, cu variația mărginită pe $[a, b]$ și pentru orice șir de diviziuni Δ_n ale lui $[a, b]$, de normă tinzînd la zero, avem

$$\lim_{n\to\infty} \omega(\Delta_n) = \overset{b}{\underset{a}{V}}(f).$$

Fie $\omega = \sum_k \omega_k$, unde $\omega_k = \omega(f; [x_k, x_{k+1}])$. Există pe $[x_k, x_{k+1}]$ două puncte ξ_k, η_k astfel încît $|f(\xi_k) - f(\eta_k)| > \omega_k - \frac{\varepsilon}{n}$, unde $\varepsilon > 0$ este dat, iar n este numărul de intervale ale diviziunii Δ. Avem deci un v dat de

$$v = \sum |f(\xi_k) - f(\eta_k)| + \sum |f(\xi_{k-1}) - f(\xi_k)| > \sum \omega_k - \varepsilon = \omega - \varepsilon.$$

Deci, $\sup_\Delta \omega(\Delta) \leqslant \sup_\Delta v(\Delta)$. Însă, în orice caz,

$$\sup_\Delta v(\Delta) \leqslant \sup_\Delta \omega(\Delta).$$

Rezultă deci că $\overset{b}{\underset{a}{V}}(f) = \sup \omega(\Delta)$. Se aplică acum raționamentul prin care se arată, în teoria integralei, că pentru orice șir de normă tinzînd la zero avem, notînd cu s suma Darboux inferioară,

$$\lim_{n \to \infty} s(\Delta_n) = \int_a^b f(x)\, dx,$$

unde se va înlocui $s(\Delta)$ cu $\omega(\Delta)$, iar $\int_a^b f(x)dx$ prin $\overset{b}{\underset{a}{V}}(f)$.

Deci, pentru orice f, cu variație mărginită pe $[a, b]$, avem

$$\lim_{\mathcal{F}^*} \omega = \overset{b}{\underset{a}{V}}(f).$$

INTEGRALELE LUI DARBOUX ALE UNEI FUNCȚII MĂRGINITE

Să notăm cu $s(\Delta)$ suma Darboux inferioară asociată, pentru o funcție f, mărginită pe $[a, b]$, unei diviziuni Δ. Avem, prin definiție,

$$\sup_\Delta s(\Delta) = \int_a^b f(x)\, dx.$$

Observăm că se obține

$$\lim_{\mathcal{F}} s = \int_a^b f(x)\,dx,$$

unde $\mathcal{F}$ este filtrul folosit pentru exprimarea variației totale.

Este suficient pentru aceasta să observăm că rămîn valabile raționamentele pentru variația totală. În locul funcției v se consideră funcția s, definită pe aceeași mulțime Δ.

Cu mici modificări în raționamente, se obține și

$$\lim_{\mathcal{F}} S = \int_a^b f(x)\,dx,$$

unde $S(\Delta)$ este suma Darboux superioară.

Spre deosebire însă de variația totală, în a cărei expresie nu se putea înlocui, în general, filtrul $\mathcal{F}$ prin filtrul $\mathcal{F}^*$, aici este posibil acest lucru, deoarece pentru $\varepsilon > 0$ există $\eta_\varepsilon > 0$ astfel încît din $\| \Delta \| < \eta_\varepsilon$ să rezulte

$$\int_a^b (f(x)\,dx - s(\Delta) < \varepsilon, \quad S(\Delta) - \int_a^{\bar{b}} f(x)dx < \varepsilon.$$

Avem

$$\lim_{\mathcal{F}^*} s = \int_a^b f(x)dx, \quad \lim_{\mathcal{F}^*} s = \int_a^{\bar{b}} f(x)dx.$$

Ținînd seama că pentru o funcție integrabilă Riemann pe $[a, b]$ cele două integrale Darboux sînt egale, avem, punînd

$$\tau(\Delta) = \sum f(\xi_i)(x_{i+1} - x_i), \quad x_i \leqslant \xi_i \leqslant x_{i+1},$$

$$\lim_{\mathcal{F}^*} s = \lim_{\mathcal{F}^*} S = \lim_{\mathcal{F}} s = \lim_{\mathcal{F}} S = \int_a^b f(x)dx =$$

$$\lim_{\mathcal{F}^*} \tau = \lim_{\mathcal{F}} \tau,$$

ori de cîte ori funcția f este integrabilă Riemann pe $[a, b]$.

În particular, rezultă că și procesul de trecere la limită care intervine în definiția integralei constituie o limită după un anumit filtru.

III

Ce este lungimea unei curbe?

INTRODUCERE

În capitolul de față încercăm să expunem evoluția noțiunii de lungime a unei curbe.

Pornind de la conținutul intuitiv al noțiunilor de „curbă" și „lungime", oamenii de știință au căutat de multă vreme să găsească cea mai potrivită traducere matematică a lui. În ciuda aparenței de simplitate pe care o prezintă această problemă, ea s-a dovedit a ascunde dificultăți și o teorie matematică completă a lungimii curbelor nu a putut fi construită decît în secolul nostru. Efortul pentru crearea acestei teorii a dat un impuls puternic dezvoltării unor importante capitole de matematică, de pildă teoria integralei, și chiar a unei întregi ramuri a matematicii, analiza funcțională.

În ce privește teoria ariei unei suprafețe, despre care o bună bucată de vreme matematicienii aveau impresia — greșită — că poate fi desfășurată ca o simplă transpunere a considerațiilor clasice asupra lungimii curbelor, trebuie să spunem că această teorie se desăvîrșește abia în zilele noastre.

În teoria lungimii și ariei rezultatele fundamentale au la bază observații atît de simple — și totuși atît de profunde —, observații legate de faptul banal de măsurare a lungimii unui drum sau a unui fir, încît cititorul rămîne oarecum surprins că matematicienii au putut să treacă atîta timp pe lîngă ele fără să le observe sau să le observe fără a trage consecințele din această observare.

Pentru a nu pierde din vedere linia principală, evoluția ideilor, vom înlătura, pe cît este posibil, detaliile și demonstrațiile.

O EXPUNERE CLASICĂ A NOȚIUNII DE LUNGIME A CERCULUI

În definiția lungimii unei curbe apar două feluri de dificultăți. Unele provin din însăși complexitatea noțiunii de curbă, noțiune, care depășește intuiția obișnuită despre linie sau despre o trăsătură continuă; altele provin din netranspunerea fidelă în matematică a conținutului intuitiv al noțiunii de lungime.

Pentru a reține doar dificultățile de al doilea tip, este suficient să considerăm o curbă de o structură mai simplă, de pildă, un cerc.

O expunere clasică a noțiunii de lungime a unei circumferințe se găsește, de exemplu, în cartea lui Jacques Hadamard, *Lecții de geometrie elementară*. Editura Tehnică, București, 1960. Această expunere se poate rezuma astfel: se consideră un șir de poligoane regulate înscrise în circumferință și un șir de poligoane regulate circumscrise circumferinței, astfel încît fiecare poligon să aibă un număr dublu de laturi față de poligonul precedent. Poligoanele inițiale din cele două șiruri au același număr de laturi. Lungimile poligoanelor din primul șir merg crescînd, ale celor din al doilea șir merg descrescînd. Orice poligon din primul șir are lungimea mai mică decît orice poligon din al doilea șir.

Pe baza teoremei: „orice șir monoton și mărginit are o limită (finită)", se deduce că șirul lungimilor poligoanelor circumscrise și șirul lungimilor poligoanelor înscrise au fiecare o limită. Se arată că raportul perimetrelor a două poligoane de același rang din cele două șiruri tinde la 1 și se stabilește astfel egalitatea celor două limite. În sfîrșit, se demonstrează că valoarea acestei limite nu depinde nici de șirul particular de poligoane înscrise sau circumscrise, nici de faptul că aceste poligoane sînt sau nu regulate sau de faptul că numărul laturilor se dublează de fiecare dată; esențial este doar ca lungimea celei mai mari laturi a poligonului (înscris sau circumscris) să tindă la zero atunci cînd rangul tinde la infinit.

Care este punctul de pornire în încercarea de a defini lungimea cercului? Deoarece știm ce înseamnă lungimea unui segment de dreaptă (aceasta presupune cunoașterea noțiunii de număr real), deci a unei linii poligonale, este natural să încercăm să definim lungimea cercului cu ajutorul lungimilor liniilor poligonale care se apropie din ce în ce mai mult de cerc. Trebuie deci să găsim traducerea matematică a procesului reprezentat prin cuvintele „se apropie". Pare firească următoarea definiție:

O linie poligonală P este la distanță inferioară lui $\varepsilon > 0$ față de circumferința C de rază r, dacă P este situată într-o coroană circulară concentrică cu C, ale cărei cercuri frontieră au razele $r - \varepsilon$ și $r + \varepsilon$.

Cîteva exemple care oglindesc această situație se pot vedea în fig. 1.III.

Dintre aceste exemple, reține în special atenția ultimul. Apropierea unei linii poligonale de o circumferință nu este condiționată de faptul că această linie poligonală este înscrisă sau circumscrisă circumferinței. Considerarea, cu exclusivitate, a poligoanelor înscrise și a celor circumscrise apare astfel arbitrară.

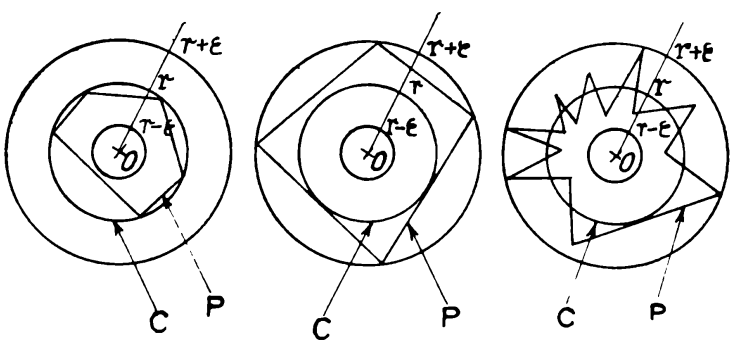


Fig. 1.III

Sîntem tentați să introducem următoarea definiție: lungimea circumferinței este limita către care tinde șirul lungimilor liniilor poligonale P_1, P_2, ..., P_n ,... alese

în așa fel, încît pentru orice n, distanța de la P la circumferință să fie inferioară lui $\frac{1}{n}$ (în sensul definiției de mai sus).

Deși luarea în considerație a liniilor poligonale de orice fel este un mare pas înainte, trebuie totuși să renunțăm la această definiție. În adevăr, dacă am adopta-o, am putea demonstra că lungimea oricărei circumferințe este egală cu infinit. Este suficient să observăm că oricît de mic ar fi $\varepsilon > 0$ și oricît de mare ar fi numărul A, există o linie poligonală P, situată față de circumferință la o distanță inferioară lui ε și astfel încît lungimea lui P să fie superioară lui A. Aceasta apare clar în fig. 2.III. Dacă unghiul dintre două laturi consecutive ale liniei poligonale este suficient de mic, atunci numărul laturilor din care este formată linia P va fi destul de mare pentru ca lungimea lui P să întreacă valoarea A.

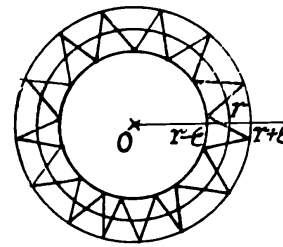


Fig. 2.III

Cel care a știut să utilizeze în mod judicios liniile poligonale apropiate de o curbă, în vederea unei definiții cît mai naturale a lungimii unei curbe și care să nu contrazică — ca cea de mai sus — nici un fapt intuitiv, a fost matematicianul Henri Lebesgue. Nu este, poate, lipsit de interes să spunem că Lebesgue și-a construit teoria sa pornind de la lucruri și observații elementare, mărturisind că această teorie are ca punct de plecare nemulțumirile pe care le avea ca elev de liceu în legătură cu modul în care se prezentau în manualele școlare, noțiunile de lungime și de arie. Aceste nemulțumiri vizau lucruri pe cît de simple, pe atît de profunde, legate mai întîi de lungimea cercului, aria cilindrului, a conului și, mai tîrziu, de teoria lui Camille Jordan, care domina la sfîrșitul secolului trecut concepțiile matematicienilor asupra lungimii și ariei. La început, Lebesgue a considerat satisfăcătoare teoria lui Jordan, dar ulterior, văzînd la ce încurcătură poate ea să conducă chiar într-un caz simplu ca aria unui cilindru de rotație, n-a mai fost mulțumit de această teorie.

Lebesgue și-a expus ideile sub forma unei critici aduse teoriei lui Jordan. Ele vor face obiectul unor paragrafe ulterioare ale acestui capitol. Să vedem deocamdată în ce constau ideile lui Jordan. Cu aceasta trecem, implicit, la discutarea acelor dificultăți care apar în definiția lungimii unei curbe, fiind provocate de însăși complexitatea noțiunii de curbă.

DIFICULTĂȚI ȘI POTICNIRI ÎN ÎNCERCĂRILE DE DEFINIRE A NOȚIUNII DE CURBĂ

După Jordan, o curbă [1] este o mulțime de puncte ale căror coordonate x, y sînt date de o reprezentare parametrică

$$x = f(t), \qquad y = g(t), \tag{1}$$

funcțiile f și g fiind continue pe un același segment $[a, b]$. Este ușor de văzut că curbele obișnuite în geometria analitică, de exemplu, sînt curbe și în sensul definiției lui Jordan. Astfel, reprezentările parametrice

$$\begin{cases} x = t \\ y = t \end{cases}, \qquad \begin{cases} x = r\cos t \\ y = r\sin t \end{cases}, \qquad \begin{cases} x = a\cos t \\ y = b\sin t \end{cases}, \tag{2}$$

unde $t \in [0, 2\pi]$, definesc respectiv un segment de dreaptă, un cerc de rază r cu centrul în originea axelor și elipsa de semiaxe a, b, centrată de asemenea în originea axelor.

Există însă curbe jordaniene care trec mult dincolo de intuiția pe care o avem despre o curbă. Astfel, Peano a dat un exemplu de curbă jordaniană care umple un pătrat. După Peano, mulți alți matematicieni au dat exemple similare. Unul dintre ele este reprodus în volumul II de *Analiză matematică*, apărut în Editura Academiei, al acad. Miron Nicolescu (p.14). Aceasta înseamnă că pătratul plin este și el o curbă, lucru care desigur ne face să zîmbim. Acest fapt curios se explică prin aceea că liniile cu care sîntem familiarizați, și anume cele folosite

[1] Ne restrîngem la curbe plane.

în geometrie, sînt definite într-un mod foarte restrictiv, cerîndu-se ca funcțiile care intervin în reprezentarea lor parametrică să aibă și alte proprietăți, în afară de continuitate. Este un fapt cunoscut că graficul unei funcții continue nu este, în general, susceptibil de un desen. Trasarea unei linii pe o foaie de hîrtie presupune că imprimăm mîinii în fiecare moment o direcție determinată. Deci linia trasată va avea tangentă în fiecare punct (exceptînd eventual un număr finit de puncte).

În afară de aceasta, trebuie să menționăm că nici definiția curbei jordaniene nu este completă. O curbă nu este doar o mulțime de puncte definită de (1), ci o astfel de mulțime, împreună cu reprezentarea ei parametrică, adică mulțimea de puncte date de (1), împreună cu reprezentarea (1).

Aici însă apar unele complicații, deoarece, cu precizarea adusă definiției noțiunii de curbă, se poate întîmpla ca două reprezentări care coincid vizual, avînd același grafic, să definească totuși curbe distincte. De pildă, reprezentările parametrice

$$\begin{cases} x = t \\ y = 2t \end{cases}, \qquad \begin{cases} x = t^2 \\ y = 2t^2 \end{cases}, \qquad (0 \leqslant t \leqslant 1) \qquad (3)$$

definesc aceeași mulțime Γ de puncte, și anume segmentul de dreaptă care unește originea axelor cu punctul (1, 2). La prima vedere, sîntem tentați să spunem că acest segment reprezintă două curbe diferite, după cum este considerat cu prima sau cu a doua reprezentare parametrică. Însă cele două reprezentări induc aceeași relație de ordine pe Γ, în sensul că dacă pentru $P_1 \in \Gamma$, $P_2 \in \Gamma$ avem $P_1 > P_2$ (adică valoarea t_1 a parametrului, căruia îi corespunde P_1, este mai mare decît valoarea t_2, căreia îi corespunde punctul P_2) după prima reprezentare, atunci avem $P_1 > P_2$ și după a doua reprezentare și reciproc. Relația de ordine este lipsită de ambiguitate, neexistînd nici un punct de pe Γ care să corespundă la două valori distincte ale lui t. După cum se vede, nu ar fi firesc să spunem că avem aici două curbe diferite.

Se pot însă ivi și cazuri mai complicate. Fie, de pildă, o funcție $f(t)$, definită și derivabilă pe [0, 1], avînd ca mul-

țime a valorilor segmentul [0, 1] și astfel încît derivata $f'(t)$ să-și schimbe semnul în cel puțin un punct din [0, 1]. Mai general, s-ar putea considera $f(t)$ continuă și nemonotonă pe [0, 1], cu mulțimea valorilor dată tot de [0, 1]). Reprezentările parametrice

$$\begin{cases} x = t \\ x = 2t \end{cases}, \qquad \begin{cases} x = f(t) \\ y = 2f(t) \end{cases}, \qquad (0 \leqslant t \leqslant 1) \quad (4)$$

definesc aceeași mulțime de puncte, dar nu induc aceeași ordine pe această mulțime. Propriu-zis, a doua reprezentare nici nu induce o ordine determinată, deoarece există puncte care corespund la mai multe valori ale lui t.

De exemplu:

$$f(t) = 8t^3 - 9t^2 + 2t.$$

Punctul M de coordonate $(f(t), 2f(t))$ descrise, cînd t variază de la 0 la 1, segmentul Γ de mai sus, dar ca în

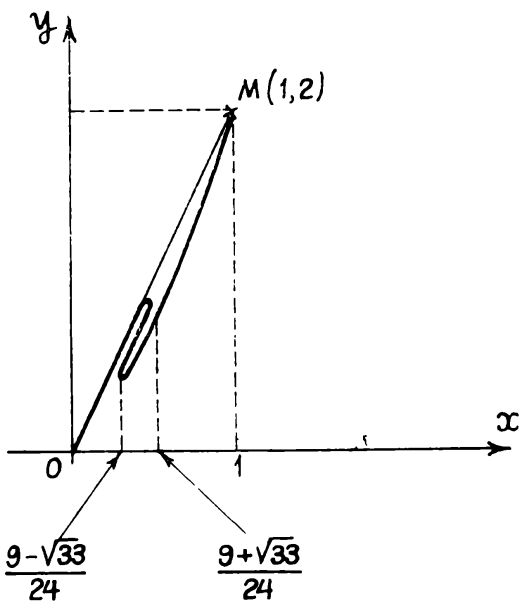


Fig. 3.III

fig. 3.III. Aceasta rezultă din faptul că o funcție continuă care nu ia nici o valoare de mai multe ori este monotonă; or, prin ipoteză, $f(t)$ nu este monotonă. Deci este legitim

să considerăm că reprezentările (4), spre deosebire de (3), definesc curbe distincte. Fenomene de acest tip au mai fost întîlnite în analiză. De pildă, două șiruri distincte de numere reale pot să parcurgă fiecare în parte aceeași mulțime, de exemplu mulțimea numerelor raționale.

Discuția de mai sus învederează necesitatea de a introduce cîteva definiții.

Un punct P al unei curbe jordaniene, dată de (1), se numește punct multiplu, dacă există cel puțin două valori t_1 și t_2 din $[a, b]$, dintre care cel puțin una este diferită atît de a, cît și de b și astfel încît $f(t_1) = f(t_2)$, $g(t_1) = g(t_2)$.

Dreapta, cercul, elipsa — cu reprezentările date de (2) — nu admit puncte multiple. În schimb, foliul lui Descartes, definit de

$$x = \frac{3t}{1 + t^3}, \qquad y = \frac{3t^2}{1 + t^3},$$

admite un punct multiplu în originea axelor (fig. 4.III).

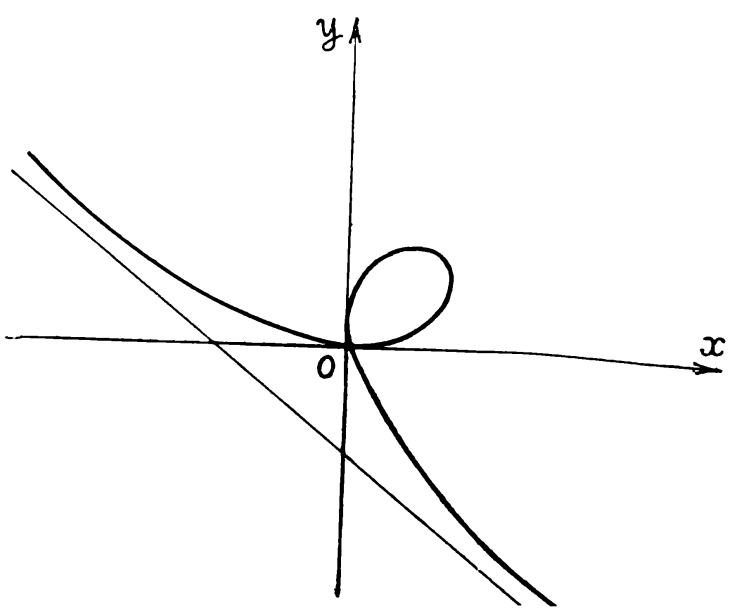


Fig. 4.III

Din cele văzute rezultă că tocmai prezența punctelor multiple complică structura unei curbe și oferă dificultăți în distingerea ei de o altă curbă, cu același grafic.

Se poate arăta, dar nu vom face aceasta aici, că punctele multiple sînt „vinovate" și de existența curbelor care umplu un pătrat, cu alte cuvinte, dacă o curbă umple un pătrat atunci ea are puncte multiple.

O curbă jordaniană care nu admite nici un punct multiplu se numește curbă simplă. Dacă $f(a) = f(b)$, $g(a) = g(b)$, atunci avem o curbă închisă. O curbă simplă care nu este închisă se numește un arc simplu. Punctele $A(f(a), g(a))$ și $B(f(b), g(b))$ se numesc, în acest caz, extremitățile arcului. Cercul și elipsa date de (2) sînt curbe simple închise. Segmentul de dreaptă dat de prima reprezentare din (2) este un arc simplu.

NOȚIUNEA DE CURBĂ ÎN ANALIZĂ

Discuția de mai sus conduce în mod natural la cîteva definiții noi, care oglindesc cu mai multă fidelitate, într-un mod mai nuanțat, complexitatea conceptului intuitiv de curbă și care ne vor da posibilitatea să studiem în mod riguros și conceptul de lungime a unei curbe.

Numim drum orice aplicație continuă a unui segment $[a, b]$ în R^2 (adică în plan) sau în R^3 (adică în spațiu). În primul caz spunem că avem un drum plan, în al doilea, un drum în spațiu. Pentru simplitate, vom considera în mod explicit numai drumuri plane, dar considerațiile se extind fără dificultate și la drumuri în spațiu.

Fie două drumuri plane d_1 și d_2. Fie $[a, b]$ segmentul de definiție a lui d_1 și fie $[\alpha, \beta]$ segmentul de definiție a lui d_2. Vom spune că drumurile d_1 și d_2 sînt echivalente dacă există o aplicație φ definită pe $[a, b]$, cu valori în $[\alpha, \beta]$, continuă și strict crescătoare pe $[a, b]$ și astfel încît $\varphi(a) = \alpha$, $\varphi(b) = \beta$ și $d_1(t) = d_2(\varphi(t))$ pentru orice t din $[a, b]$. Relația introdusă este o veritabilă relație de echivalență; reflexivitatea se verifică luînd în rolul lui φ aplicația identică; simetria este o consecință a faptului că inversa unei funcții continue și strict crescătoare este o funcție cu aceleași proprietăți; tranzitivitatea rezultă din faptul că suprapunerea a două funcții continue și strict crescătoare este tot o funcție continuă, strict crescătoare.

În felul acesta, toate drumurile se repartizează în clase de echivalență. Fiecare dintre aceste clase va constitui, prin definiție, o curbă.

Acest mod de a introduce noțiunea de curbă poate să surprindă pe unii cititori; într-adevăr, există în matematica elementară obiceiul de a considera conceptul de curbă ca fiind de natură geometrică, în timp ce aici acest concept se definește ca o clasă de aplicații continue, deci este de natură esențial analitică. Apoi, o curbă este aici privită ca rezultat al unui proces de identificare a unor aplicații considerate indiscernabile din punctul de vedere pe care l-am adoptat, deci ca o clasă de echivalență în interiorul căreia alegerea unui element este o chestiune de comoditate, dar nu de principiu.

Caracterul analitic al conceptului de curbă nu trebuie totuși să ne împiedice să vedem componenta sa geometrică. O teoremă importantă afirmă că dacă două drumuri sînt echivalente, ele au același grafic; este suficient, pentru a obține același rezultat, să ținem seamă că atunci cînd t parcurge pe $[a, b]$, $\varphi(t)$ parcurge pe $[\alpha, \beta]$, iar cînd τ parcurge pe $[\alpha, \beta]$, $\varphi^{-1}(\tau)$ parcurge pe $[a, b]$, în timp ce pentru oricare valoare a lui t din $[a, b]$ avem egalitatea $d_1(t) = d_2(\varphi(t))$. (Am folosit notațiile din definiția relației de echivalență a două drumuri.) De aici rezultă că fiecărei curbe i se asociază, în mod natural, o anumită mulțime de puncte care constituie graficul curbei și care nu este altceva decît graficul comun tuturor drumurilor din clasa corespunzătoare de echivalență. Acest grafic constituie componenta geometrică a conceptului analitic de curbă, componentă prin intermediul căreia acest concept se înrudește cu conceptul geometric corespunzător.

Fie (x_0, y_0) un punct de pe graficul d definit de (1). Spunem că acest punct este simplu pentru d dacă nu există două valori t_1, t_2 din $[a, b]$, $t_1 \neq t_2$ (dintre care cel puțin una este diferită atît de a, cît și de b) și astfel încît să avem $d(t_1) = d(t_2) = (x_0, y_0)$. Un punct care nu este simplu se numește multiplu. Un drum fără puncte multiple se numește drum simplu.

Drumul d dat de (1) este prin definiție închis dacă avem $d(a) = d(b)$.

Un număr de teoreme relativ simple ne asigură că două drumuri echivalente sînt de aceeași natură, în sensul că dacă unul dintre ele este simplu (sau închis), atunci și celălalt este simplu (respectiv închis). Aceasta ne permite să putem vorbi despre curbe simple (deci care conțin numai drumuri simple) și despre curbe închise (deci care conțin numai drumuri închise).

Dacă însă fiecărei curbe i se asociază în mod natural o componentă geometrică, nu trebuie să ne scape faptul important că această componentă geometrică *nu* caracterizează curba, *nu* o individualizează. Într-adevăr, este posibil ca două curbe distincte să aibă același grafic. Astfel, drumurile date de reprezentările (4) nu sînt echivalente, deoarece primul este simplu, pe cînd al doilea nu este simplu. Deci aceste două drumuri definesc două curbe distincte. Totuși, aceste curbe au același grafic și anume segmentul de dreaptă care unește originea axelor cu punctul de abscisă 1 și ordonată 2. În schimb, cele două drumuri definite de reprezentările (3) sînt echivalente, după cum se poate verifica luînd în rolul funcției φ, funcția $\varphi(t) = t^2$. Deci aceste două drumuri definesc una și aceeași curbă.

Această concepție despre curbă, privită ca o clasă de drumuri echivalente, ne amintește de un alt concept important al analizei matematice, acela de număr real. Noțiunea de număr real se definește de asemenea ca o clasă de echivalență, anume ca o clasă de șiruri convergente de numere raționale, echivalența a două șiruri revenind la faptul că șirul diferență converge către zero. Ne amintim că în practica obișnuită a analizei matematice numărul real este identificat cu un șir din clasa care-l definește, lucru permis datorită faptului că proprietățile unui număr real sînt implicate în oricare dintre șirurile care-l reprezintă. Dar această situație reclamă o deosebită atenție, deoarece o noțiune sau operație relativă la șiruri convergente de numere raționale nu poate deveni o noțiune sau operație relativă la numere reale decît după ce se demonstrează invarianța ei față de relația de echivalență a două șiruri convergente de numere raționale. Astfel, ne amintim, pentru ca operația de adunare a două astfel de șiruri să poată fi ridicată la rangul de operație de adu-

nare a două numere reale era necesar să demonstrăm, în prealabil, că rezultatul acestei operații nu se schimbă atunci cînd unul dintre șiruri este înlocuit cu un altul, echivalent cu el.

Precauții asemănătoare se impun și în studiul curbelor. Putem identifica, pentru comoditate, o curbă cu unul oarecare dintre drumurile care o definesc, putem considera aceste noțiuni ca fiind indiscernabile, cu condiția de a nu uita, oridecîteori se introduce o noțiune sau operație relativă la drumuri, că ele nu pot deveni apanajul curbelor corespunzătoare decît dacă se demonstrează invarianța lor față de relația de echivalență a două drumuri.

CE ESTE UN DRUM RECTIFICABIL ?

Fie d un drum dat de (1). Fie $\Delta = (a = t_0 < t_1 < \ldots$ $\ldots < t_i < t_{i+1} < \ldots < t_n = b)$ o diviziune a segmentului $[a, b]$. Fie P_i punctul de coordonate $f(t_i), g(t_i)$. Fie Π_Δ linia poligonală formată prin unirea, două cîte două, în ordinea scrisă, a punctelor P_0, P_1, P_2, ..., P_i, P_{i+1}, ..., P_n. Π_Δ este înscrisă în d, în sensul că vîrfurile lui Π_Δ sînt situate pe graficul lui d. Fie $l(\Pi_\Delta)$ lungimea lui Π_Δ. Spunem că d este rectificabil în sensul lui Jordan, dacă mulțimea numerelor $\{l(\Pi_\Delta)\}$ este mărginită atunci cînd Δ parcurge toate diviziunile posibile ale lui $[a, b]$. Cel mai mic număr K (se demonstrează existența lui!), astfel încît pentru orice Δ, $l(\Pi_\Delta) \leqslant K$, se numește lungimea drumului d. Notăm această lungime cu $l(d)$.

Definiția dată comportă unele discuții. Mai întîi, poate să pară curios că nu considerăm totalitatea liniilor poligonale înscrise în d, ci doar acele linii poligonale care sînt induse de diviziuni ale lui $[a, b]$. Dacă însă nu am proceda așa, definiția pusă n-ar mai fi bună la nimic, pentru că ar fi lipsită de obiect. N-ar mai exista nici un drum rectificabil; fiecare drum ar avea o lungime infinită. Să luăm, de pildă, o circumferință (fig. 5. III). Este vizibil că putem înscrie în ea o linie poligonală a cărei lungime să întreacă un număr dat, unind, două cîte două, diverse puncte ale circumferinței de un număr suficient de mare de ori.

Se pune în mod firesc următoarea întrebare: definiția elementară a lungimii cercului (considerat cu reprezentarea parametrică din (2)) coincide cu definiția dată mai sus, datorită lui Jordan? Răspunsul este afirmativ. Pentru a-l obține, vom da mai întîi un rezultat general.

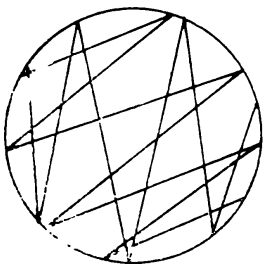


Fig. 5.III

Fie $\Delta = (a = t_0 < t_1 < \ldots < t_i < t_{i+1} < \ldots < t_n = b)$ o diviziune a lui $[a, b]$. Numărul

$$\nu(\Delta) = \max_{0 \leqslant i \leqslant n-1} \{(t_{i+1} - t_i)\}$$

se numește norma diviziunii Δ. Fiecărei diviziuni Δ a lui $[a, b]$ îi corespunde o linie poligonală Π_Δ, înscrisă în d și având drept vîrfuri punctele $P_i(f(t_i)), g(t_i))$; $(i = 0, \ldots, n)$. Se poate arăta că pentru un șir de diviziuni Δ_n, pentru care

$$\lim_{n \to \infty} \nu(\Delta_n) = 0$$

avem:

$$\lim_{n \to \infty} l(\Pi_{\Delta_n}) = l(d).$$

Pentru demonstrarea acestei afirmații se folosește următoarea observație: Dacă orice punct de diviziune în Δ' este punct de diviziune în Δ'', atunci $l(\Pi_{\Delta'}) \leqslant l(\Pi_{\Delta''})$. Nu vom da această demonstrație.

Deoarece drumul cu care lucrăm este simplu, se poate arăta că dacă $\lim \nu(\Delta_n) = 0$, atunci și lungimea celei mai mari laturi a liniei poligonale Π_{Δ_n} tinde la zero și reciproc. De aici rezultă: lungimea, în sensul lui Jordan, a unui drum simplu, poate fi privită ca limita lungimilor unor linii poligonale înscrise în drum, pentru care lungimea celei mai mari laturi tinde la zero. (Lucrăm tot timpul numai cu linii poligonale induse de diviziuni ale lui $[a, b]$, prin intermediul reprezentării (1).)

Să ne întoarcem acum la lungimea cercului. Un poligon regulat Π_n cu n laturi, înscrise într-un cerc, corespunde unei anumite diviziuni Δ_n a intervalului $[0, 2\pi]$, diviziune formată cu acele valori ale lui t cărora le cores-

pund, prin a doua reprezentare din (2), vîrfurile poligonului Π_n. Atunci cînd $n \to \infty$, lungimea unei laturi a poligonului regulat înscris Π_n, tinde la zero, deci $\lim_{n\to\infty} \nu(\Delta_n) = 0$. De aici rezultă, conform propoziției generale de mai sus, că

$$\text{lungimea cercului} = \lim_{n\to\infty} l(\Pi_n),$$

lungimea fiind luată în sensul definiției lui Jordan. Prin aceasta s-a demonstrat, implicit, că cercul (considerat cu reprezentarea din (2)) este un drum rectificabil.

UN CRITERIU DE RECTIFICABILITATE

Fie un drum d dat de reprezentarea (1).

Este important să se stabilească ce proprietăți trebuie să aibă aceste funcții pentru ca d să fie rectificabil și, dacă este posibil, să se exprime însăși lungimea drumului cu ajutorul acestor funcții. Soluția primei probleme o vom da imediat; soluția celei de-a doua va fi dată ceva mai tîrziu.

Mai întîi să amintim o noțiune din §11, capitolul II. O funcție reală $f(t)$, definită pe $[a, b]$, este cu variație mărginită pe $[a, b]$, dacă există un număr K, astfel încît, oricare ar fi diviziunea $\Delta = (a = t_0 < t_1 < \ldots < t_i < t_{i+1} < \ldots < t_n = b)$, avem:

$$v(\Delta) = \sum_{i=0}^{n-1} |f(t_{i+1}) - f(t_i)| \leqslant K. \qquad (5)$$

Pentru a înțelege sensul intuitiv al acestei noțiuni, este suficient să observăm semnificația expresiei lui $v(\Delta)$, dată de (5). Numărul $v(\Delta)$ însumează toate modificările pe care le suferă funcția pe parcursul de la a la b, dacă se face abstracție de valorile pe care ea le ia în interiorul intervalelor care alcătuiesc diviziunea Δ. Este clar atunci că pentru a avea o măsură cît mai bună a modificărilor pe care valorile funcției le suferă între a și b este indicat să îndesim cît mai mult punctele diviziunii Δ. Cum numărul $v(\Delta)$ nu scade atunci cînd adău-

găm noi puncte diviziunii Δ, este natural să se îndrepte atenţia asupra celui mai mic număr K cu proprietatea (5). (Se poate arăta că un astfel de număr există.) Acest număr va da cea mai bună măsură a modificărilor pe care funcţia $f(x)$ le suferă pe segmentul $[a, b]$. Cel mai mic număr K cu proprietatea (5) se numeşte variaţia totală a funcţiei $f(t)$ de la a la b şi se notează $\overset{b}{\underset{a}{V}}(f)$ (a se vedea şi § 11 din capitolul II).

O funcţie cu variaţie mărginită pe $[a, b]$ este tot una cu o funcţie care se poate scrie ca diferenţă a două funcţii monoton crescătoare pe $[a, b]$. Într-adevăr, dacă notăm cu $\varphi(x)$ variaţia totală de la a la x a funcţiei f şi cu $\psi(x)$ diferenţa $\varphi(x) - f(x)$, funcţiile φ şi ψ sînt monoton crescătoare pe $[a, b]$. Orice funcţie $f(t)$ cu derivată mărginită pe $[a, b]$ este cu variaţie mărginită pe $[a, b]$. În adevăr, aplicînd formula creşterilor finite, avem, pentru o diviziune Δ arbitrară,

$$v(\Delta) = \sum_{i=0}^{n-1} |f(t_{i+1}) - f(t_i)| = \sum_{i=0}^{n-1} |f'(\xi_i)|\,(t_{i+1} - t_i),$$
$$(t_i < \xi_i < t_{i+1}).$$

Prin ipoteză, există un număr $M > 0$, astfel încît $|f'(t)| < M$, pentru orice $t \in [a, b]$, deci

$$\sum_{i=0}^{n-1} |f'(\xi_i)|\,(t_{i+1} - t_i) \leqslant \sum_{i=0}^{n-1} M(t_{i+1} - t_i) = M(b - a),$$

deci

$$v(\Delta) < M(b - a),$$

pentru orice diviziune Δ a lui $[a, b]$.

Iată un exemplu simplu de funcţie care nu este cu variaţie mărginită (fig. 6.III):

$$f(x) = \begin{cases} \sin \dfrac{1}{x}, & x \neq 0 \\ 0, & x = 0 \end{cases} \qquad (-1 \leqslant x \leqslant 1).$$

Pe segmentul $[-1, 1]$ există o infinitate de puncte x în care $f(x) = 0$ $\left(\text{anume } x_n = \dfrac{1}{n\pi}\right)$ şi o infinitate de puncte x în care $f(x) = \pm 1$ $\left(\text{anume } x'_n = \dfrac{2}{(2n+1)\pi}\right)$. Lu-

înd un număr destul de mare de puncte x_n și x'_n, vom putea forma cu aceste puncte o diviziune Δ, pentru care $v(\Delta)$ să fie oricît de mare dorim.

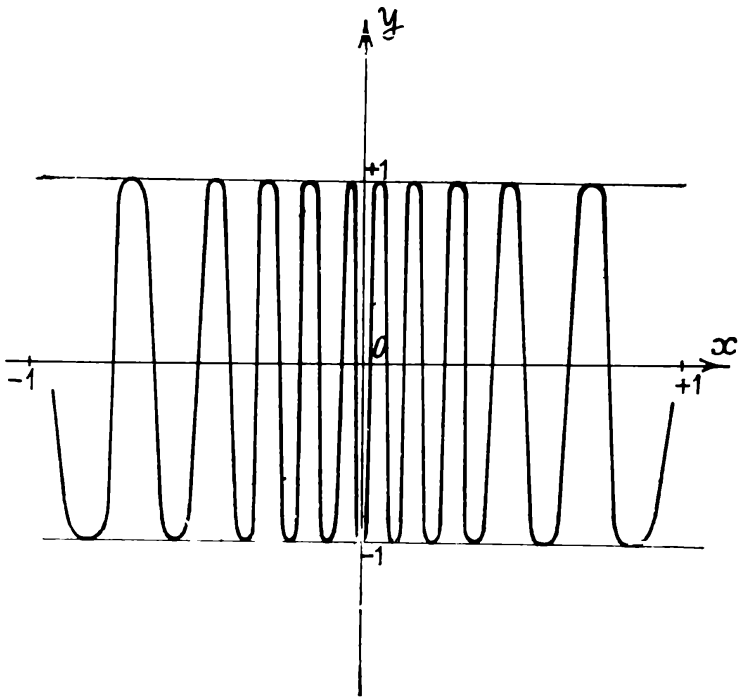


Fig. 6.III

Vom demonstra acum următoarea teoremă a lui Jordan: necesar și suficient ca drumul d dat de (1) să fie rectificabil, este ca funcțiile f și g să fie cu variație mărginită pe $[a, b]$.

Într-adevăr pentru o diviziune Δ a lui $[a, b]$ avem:

$$\left.\begin{aligned}\sum_{i=0}^{n-1}|f(t_{i+1})-f(t_i)|\\ \sum_{i=0}^{n-1}|g(t_{i+1})-g(t_i)|\end{aligned}\right\}\leqslant\sum_{i=0}^{n-1}\sqrt{[f(t_{i+1})-f(t_i)]^2+[g(t_{i+1})-g(t_i)]^2}\leqslant$$

$$\leqslant\sum_{i=0}^{n-1}\left(|f(t_{i+1})-f(t_i)|+|g(t_{i+1})-g(t_i)|\right), \qquad (6)$$

deoarece se știe din algebra elementară că dacă $a_1, a_2, \ldots, \ldots, a_n, b_1, b_2, \ldots, b_n$ sînt $2n$ cantități nenegative, atunci

$$\left.\begin{matrix}\sum_{i=1}^{n} a_i \\ \sum_{i=1}^{n} b_i\end{matrix}\right\} \leqslant \sum_{i=1}^{n} \sqrt{a_i^2 + b_i^2} \leqslant \sum_{i=1}^{n} (a_i + b_i).$$

Suma de la mijloc reprezintă lungimea unei linii poligonale înscrisă în d, și anume a liniei poligonale care are drept vîrfuri punctele corespunzătoare prin (1) valorilor t_i ale diviziunii Δ. Deci, dacă suma de la mijloc este, pentru orice Δ, inferioară unui K fix, atunci d este rectificabil.

Să presupunem acum că f și g sînt cu variație mărginită. Suma din dreapta va fi, pentru orice Δ, inferioară unui număr K (independent de Δ), deci cu atît mai mult va avea această proprietate suma de la mijloc. Dar aceasta înseamnă, după cum am văzut, tocmai rectificabilitatea lui d.

Reciproc, dacă d este rectificabil, atunci suma de la mijloc este, pentru orice Δ, inferioară unui K independent de Δ. Cu atît mai mult vor fi inferioare lui K cele două sume din stînga. Aceasta înseamnă că f și g sînt cu variație mărginită pe $[a, b]$.

Funcțiile care definesc în (2) reprezentarea parametrică a cercului, a elipsei sau a unui segment de dreaptă sînt cu variație mărginită, deoarece au derivată mărginită. Deci segmentul de dreaptă, cercul, elipsa date de (2) sînt curbe rectificabile. Acum este ușor să construim și un exemplu de drum simplu nerectificabil. Fie, în adevăr, funcția

$$f(t) = \begin{cases} t \sin \frac{1}{t} & t \neq 0, \\ 0, & t = 0. \end{cases}$$

Această funcție este continuă pe $[0, 1]$, dar nu este cu variație mărginită pe $[0, 1]$, cum se poate vedea ușor. Deci drumul definit de $x = f(t)$, $y = t$ $(0 \leqslant t \leqslant 1)$ este

un drum simplu, nerectificabil (absența punctelor multiple este evidentă).

Să dăm și un exemplu de două drumuri cu același grafic, unul rectificabil, iar celălalt nerectificabil. Fie d_1 și d_2 drumurile date de

$$d_1: \begin{cases} x_1 = t, \\ y = t, \end{cases} \qquad d_2: \begin{cases} x = f(t), \\ y = f(t), \end{cases} \qquad (0 \leqslant t \leqslant 1),$$

unde f este o funcție continuă care nu este cu o variație mărginită pe $[0, 1]$, iar $f(0) = 0$, $f(1) = 1$ și $0 \leqslant f(t) \leqslant 1$ pentru $0 \leqslant t \leqslant 1$. Atît d_1 cît și d_2 au ca grafic segmentul de dreaptă care unește originea axelor cu punctul $(1, 1)$; d_1 este simplu și rectificabil, deoarece $\varphi(t) = t$ este cu variație mărginită ; d_2 nu este rectificabil, deoarece $f(t)$ nu este cu variație mărginită.

Un exemplu de două drumuri cu același grafic, amîndouă rectificabile, dar avînd lungimi diferite, poate fi găsit în tratatul *Analiză matematică* al acad. Miron Nicolescu (Editura Academiei, vol. II, 1953, pp. 67—68).

NOȚIUNEA DE CURBĂ RECTIFICABILĂ

Pentru a introduce această noțiune, trebuie mai întîi să ne convingem că noțiunea de drum rectificabil este invariantă prin relația de echivalență a două drumuri. Fie un drum d dat de (1) și un drum d_1 dat de reprezentarea $x = f_1(\tau)$, $y = g_1(\tau)$, unde f_1 și g_1 sînt continue pe $[\alpha, \beta]$. Să presupunem că drumurile d și d_1 sînt echivalente și că drumul d este rectificabil, și să arătăm că drumul d_1 este de asemenea rectificabil. Există deci o funcție φ, continuă și strict crescătoare pe $[a, b]$, cu valori în $[\alpha, \beta]$, astfel încît $\varphi(a) = \alpha$, $\varphi(b) = \beta$ și $d(t) = d_1(\varphi(t))$ pentru orice t din $[a, b]$. Fiind dată o diviziune $\Delta = (a = t_0 < t_1 < \ldots < t_i < t_{i+1} < \ldots < t_n = b)$ a intervalului $[a, b]$, ei îi corespunde, prin aplicația φ, o diviziune $\Delta_1 = (\alpha = \tau_0 < \tau_1 < \ldots < \tau_i < \tau_{i+1} < \ldots < \tau_n = \beta)$ a intervalului $[\alpha, \beta]$ și reciproc, diviziunii Δ_1 îi corespunde

prin inversa funcției φ, diviziunea Δ. (Aici, $\tau_i = \varphi(t_i)$ pentru $0 \leqslant i \leqslant n$.) Avem:

$$l_\Delta(d) = \sum_{i=0}^{n} \sqrt{[f(t_{i+1}) - f(t_i)]^2 + [g(t_{i+1}) - g(t_i)]^2} =$$

$$= \sum_{i=0}^{n} \sqrt{[f_1(\varphi(t_{i+1})) - f_1(\varphi(t_i))]^2 + [g_1(\varphi(t_{i+1})) - g_1(\varphi(t_i))]^2} =$$

$$= \sum_{i=0}^{n} \sqrt{[f_1(\tau_{i+1}) - f_1(\tau_i)]^2 + [g_1(\tau_{i+1}) - g_1(\tau_i)]^2} = l_{\Delta_1}(d_1),$$

deci

$$l(d) = \sup_{\Delta} l_\Delta(d) = \sup_{\Delta_1} l_{\Delta_1}(d_1) = l(d_1)$$

și cele două drumuri au aceeași lungime. Drumul d fiind însă rectificabil, $l(d)$ este finită , deci $l(d_1)$ este finită, iar drumul d_1 este rectificabil.

Am demonstrat astfel chiar mai mult decît ne-am propus. Am arătat că nu numai proprietatea de rectificabilitate, dar și lungimea unui drum este invariantă prin relația de echivalență a două drumuri. Această situație legitimează pe deplin următoarele două definiții:

Numim curbă rectificabilă o clasă de drumuri rectificabile echivalente.

Lungimea unei curbe este, prin definiție, lungimea comună a drumurilor care definesc curba considerată.

Rezultă, în baza criteriului de rectificabilitate al lui Jordan, că dacă un drum este definit de funcții cu variație mărginită, atunci orice drum echivalent cu el este definit tot de funcții cu variație mărginită.

Cele stabilite ne permit să studiem o curbă rectificabilă prin intermediul unuia singur dintre drumurile care o reprezintă. Așadar, vom folosi uneori termenul de curbă chiar dacă, în mod efectiv, lucrăm doar cu unul dintre drumurile care îi aparține.

A sosit momentul să expunem critica făcută de Lebesgue teoriei clasice a lungimii unei curbe.

Confuzia cea mai gravă pe care teoria clasică o strecoară în capul cititorului (chiar dacă aceasta nu se face în mod explicit) este ideea că lungimea ar fi o funcție continuă de curbă. Iată cum se ajunge aici la noțiunea de continuitate:

Maurice Fréchet asociază fiecărei perechi C, Γ de curbe un număr nenegativ (C, Γ) pe care-l numește *ecartul* acelor două curbe. Acest număr este astfel, încît $(C, \Gamma) = 0$ dacă și numai dacă C și Γ coincid, $(C, \Gamma) = (\Gamma, C)$ și, fiind date trei curbe C_1, C_2 și C_3, avem $(C_1, C_2) \leqslant (C_1, C_3) + (C_2, C_3)$. Iată cum se poate defini un astfel de număr.

Fie $x = f(t)$, $y = g(t)$ și $x = \Phi(t)$, $y = \Psi(t)$, $(0 \leqslant t \leqslant 1)$ niște reprezentări parametrice ale curbelor C și Γ. (f, g, Φ și Ψ sînt continue pe $[0, 1]$). Să punem

$$\delta(t) = \sqrt{[f(t) - \Phi(t)]^2 + [g(t) - \Psi(t)]^2};$$

$\delta(t)$ este o funcție continuă pe $[0, 1]$, deci admite un maxim absolut d pe $[0, 1]$. Acest d depinde de reprezentările parametrice alese pentru C și Γ. Fie ω cel mai mic număr (se arată existența lui) din mulțimea de numere d, obținute pentru toate reprezentările parametrice posibile ale lui C și Γ. Acest număr ω satisface proprietățile din definiția ecartului și este numit ecartul curbelor C și Γ. După cum se vede, ecartul este pentru două curbe ceea ce este definiția obișnuită pentru două puncte.

Acum dispunem și de un nou mod de a aprecia depărtarea a două curbe.

Spunem că șirul de curbe $\{\Gamma_n\}$ tinde către curba Γ dacă distanța între Γ și Γ_n este inferioară lui $\varepsilon_n > 0$, iar șirul $\{\varepsilon_n\}$ tinde la zero.

Astfel, putem spune că un șir de curbe $\{C_n\}$ tinde către curba C, dacă ecartul (C, C_n) tinde la zero o dată cu $\frac{1}{n}$.

O condiție necesară și suficientă ca aceasta să se întîmple

este ca pentru orice reprezentare parametrică $x = f(t)$, $y = g(t)$, $(0 \leqslant t \leqslant 1)$ a lui C să se poată alege o reprezentare parametrică a lui C_n, fie ea $x = f_n(t)$, $y = g_n(t)$, $(0 \leqslant t \leqslant 1)$, astfel încît f_n să tindă uniform către f, iar g_n să tindă uniform către g pe $[0, 1]$, adică fiecărui $\varepsilon > 0$ să-i corespundă un N cu proprietatea că $|f_n(t) - f(t)| < \varepsilon$, $|g_n(t) - g(t)| < \varepsilon$, de îndată ce $n > N$ și oricare ar fi $t \in [0, 1]$.

O funcție reală $\varphi(\Gamma)$ de curba Γ [1] va fi, prin definiție, continuă dacă pentru orice șir $\{\Gamma_n\}$ de curbe, care tinde către Γ, avem:

$$\lim_{n \to \infty} \varphi(\Gamma_n) = \varphi(\Gamma).$$

Dacă un șir de linii poligonale tinde către un cerc în sensul definiției date în paragraful 6, atunci acest șir tinde către cerc și în sensul definiției date aici. Deci, poligoanele regulate folosite în definiția lungimii cercului formează un șir care tinde către cerc, atît în sensul definiției date în paragraful 6, cît și în sensul celei date aici.

Teoria obișnuită a lungimii cercului se bazează, în mod tacit, pe următorul deziderat, de fapt irealizabil: dacă un șir de linii poligonale tinde către circumferință, atunci lungimea circumferinței este însăși limita lungimilor liniilor poligonale. Apropierea liniilor poligonale de circumferință este concretizată în șirul de poligoane regulate înscrise, pentru care numărul laturilor tinde la infinit.

Afirmația că dezideratul de mai sus este irealizabil poate fi justificată. Aici, ne vom mulțumi să constatăm că fig. 7.III este elocventă în această privință. Pentru două numere pozitive date ε și K se pune în evidență posibilitatea ca linie poligonală să fie, față de circumferință, la distanță inferioară lui ε și de lungime superioară lui K.

S-ar putea obiecta că nu a fost aleasă judicios definiția distanței; dar această definiție nu poate fi pusă în mod arbitrar. O distanță, oricum ar fi definită, trebuie să sa-

[1] O funcție reală de curba Γ este o lege care asociază fiecărei curbe Γ un număr real unic determinat $\varphi(\Gamma)$.

tisfacă anumite cerințe intuitive. În particular, ar fi inadmisibil ca o linie conținută într-un manșon (coroană circulară) de lățime ε, care îmbracă circumferința, să nu fie considerată la distanță inferioară lui ε față de circumferință.

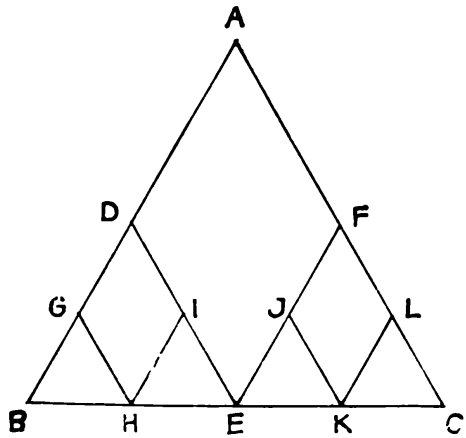


Fig. 7.III

Oricîte reproșuri am adus sau am mai aduce teoriei clasice a lungimii cercului, cititorul nu va putea fi clintit din convingerea sa că totuși definiția lungimii circumferinței cu ajutorul poligoanelor regulate înscrise conduce la un rezultat corect, adică la același număr la care ar conduce cine știe ce teorie savantă, teorie lipsită de defectele enumerate aici.

Această convingere a cititorului corespunde realității, dar ea nu face decît să mărească gravitatea greșelilor comise. Dacă un raționament greșit conduce la un rezultat absurd, însăși absurditatea rezultatului atrage atenția asupra greșelii și face imposibilă repetarea ei. Dacă însă rezultatul este corect, greșeala poate să treacă neobservată, să se repete apoi în alte probleme și să contamineze capitole întregi.

Exact așa s-a întîmplat multă vreme cu greșelile comise în definirea lungimii cercului. Aceste greșeli au trecut neobservate și matematicienii le-au repetat în definiția ariei unei suprafețe. Cît de stupefiată a rămas lumea matematică atunci cînd H. A. Schwarz a pus în evidență, printr-un exemplu devenit celebru, că nici măcar pentru o suprafață cilindrică circulară vechea definiție a ariei, transpunere fidelă a definiției lungimii cercului, nu este satisfăcătoare, putînd conduce la rezultatul absurd că un cilindru circular de rază și înălțime finite poate avea o arie infinită!

Cel mai simplu exemplu care arată că lungimea unei curbe nu este o funcție continuă de curbă este următorul:

Fie triunghiul echilateral ABC de latură 1 (fig. 7.III). Fie D, E și F mijloacele laturilor AB, BC și AC. Fie apoi G, H, I, J, K, L mijloacele segmentelor BD, BE, DE, EF, CE și CF ș.a.m.d. Linia poligonală BAC se află la distanță $\leqslant \frac{\sqrt{3}}{2}$ de BC; linia poligonală $BDEFC$ se află la distanță $\leqslant \frac{\sqrt{3}}{2\cdot 2}$ de BC; linia poligonală $BGHIEJKLC$ se află la distanță $\leqslant \frac{\sqrt{3}}{2\cdot 2^2}$ de BC; a n-a linie poligonală obținută prin procedeul arătat va fi situată la distanță $\leqslant \frac{\sqrt{3}}{2^n}$ de BC. Se obține astfel un șir de linii poligonale care tinde către segmentul BC. Printr-un raționament simplu se constată că fiecare linie poligonală din șir are lungimea constant egală cu 2. Ar însemna deci, dacă lungimea ar fi o funcție continuă de curbă, că lungimea segmentului BC să fie 2 — lucru absurd.

Prin același „raționament" s-ar putea ajunge și la alte absurdități, de pildă, egalitatea dintre lungimea unei semicircumferințe și a diametrului ei. Cu toate acestea, din păcate, astfel de procedee sînt utilizate curent în geometrie cînd se definesc lungimea cercului, aria cilindrului, a conului, a sferei.

Pentru a găsi calea justă, trebuie să ne întoarcem la originea fizică, intuitivă, a noțiunilor pe care le studiem.

Cum procedează un măsurător de terenuri cînd are de măsurat lungimea unui drum? El urmărește linia mediană sau axa drumului. Dar această linie nu poate fi determinată cu precizie; ar fi, deci, greșit să pretindem că linia poligonală, pe care el o măsoară în realitate, este înscrisă în curba — linie mediană a drumului. Din nou ajungem la constatarea că nu trebuie să bazăm definiția lungimii unei curbe pe linii poligonale înscrise. Aceasta cu atît

mai mult, cu cît o măsurătoare, oricît de bună ar fi, nu poate distinge un punct din punctele situate la o anumită apropiere de el. Deci, niciodată n-am fi în stare să precizăm dacă două măsurători distincte s-au referit sau nu la aceeași linie poligonală înscrisă.

Se creează astfel impresia că nu există nici-o posibilitate de ieșire din încurcătură. Atît considerarea liniilor poligonale înscrise, cît și a celor apropiate de curbă apar viciate.

Totuși, calea justă trebuie căutată în considerarea liniilor poligonale apropiate, vecine de curbă. Unele dintre aceste linii poligonale ne-au pus în încurcătură, deoarece pot avea lungime oricît de mare. Să facem atunci o selecție în familia șirurilor de linii poligonale care tind către curbă. Exemplul cu triunghiul echilateral, de mai sus, ne sugerează că în aproximarea unei curbe printr-o linie poligonală trebuie să evităm drumurile inutil complicate. Aceasta este, de altfel, în conformitate cu tehnica experimentală. Un măsurător de drumuri care pentru a măsura lungimea unui drum ar pune originea lanțului său pe marginea dreaptă a drumului, iar extremitatea lanțului pe marginea stîngă, apoi originea pe marginea stîngă, iar extremitatea pe marginea dreaptă ș.a.m.d. (fig. 8.III), un astfel de măsurător, spune Lebesgue, „n-ar lucra după regulile artei sale".

Or, tocmai astfel de „drumuri inutil complicate" reprezintă liniile poligonale cu care am încercat, mai sus (fig. 7.III), să aproximăm segmentul *BC*. De aici, eșecul încercării.

Problema este acum: care este instrumentul matematic cu ajutorul căruia putem împiedica șirurile de linii poligonale, inutil complicate, de a-și exercita influența lor nefastă asupra definiției lungimii unei curbe? Pentru a găsi răspunsul la această întrebare, vom face o observație simplă, care este foarte bine ilustrată în fig. 2.III: dacă un număr λ a fost obținut ca limită a unui șir de lungimi de linii poligonale tinzînd către o curbă Γ, atunci orice număr superior lui λ poate fi obținut de asemenea ca o astfel de limită. Aceasta înseamnă că tota-

litatea numerelor care pot fi obținute în acest fel se reprezintă geometric printr-o semidreaptă nemărginită la dreapta. Cu cît un număr real λ este situat mai la stînga, pe această semidreaptă, cu atît sînt mai scurte liniile poligonale care tind către curbă și ale căror lungimi tind către λ. Dacă notăm cu L extremitatea din stînga a semidreptei (evident, L este un număr pozitiv), rezultă că L este limita lungimilor celor mai scurte linii poligonale posibile, care tind către curbă. Este natural ca L să fie decretat lungimea curbei Γ. Prin definiție deci, lungimea unei curbe este cea mai mică valoare (se arată existența ei!) care se poate obține ca limită a unui șir de linii poligonale tinzînd către curbă.

Se poate întîmpla ca această „cea mai mică valoare" să fie $+\infty$. În acest caz, curba nu are lungime finită. În cazul contrar, curba se numește rectificabilă.

Această noțiune de rectificabilitate, datorită lui Lebesgue, este echivalentă cu noțiunea de rectificabilitate în sensul lui Jordan, însă îi este superioară ca metodă acesteia din urmă, fiind mai apropiată de sensul intuitiv al noțiunii de lungime și putînd fi ușor transferată la aria suprafeței. Nu vom face aici această demonstrație. Vom remarca doar că jumătate din această echivalență este evidentă. În adevăr, lungimea în sensul lui Lebesgue este vizibil mai mică sau egală cu lungimea în sensul lui Jordan.

CONVERGENȚĂ ÎN DISTANȚĂ ȘI CONVERGENȚĂ ÎN DIRECȚIE

Vom ajunge la înțelegerea acestei împrejurări fericite — echivalența celor două definiții ale lungimii — și pe o altă cale. Ne vom întoarce mult în urmă. Teoria lui Jordan a apărut la sfîrșitul secolului trecut. Nu trebuie însă să ne închipuim că pînă la Jordan matematicienii n-au fost în stare să dea o definiție noțiunii de lungime a unei curbe.

Pînă la Jordan se lucra cu definiția următoare: lungimea unei curbe, date de

$$x = f(t), \quad y = g(t), \quad (a \leqslant t \leqslant b),$$

este

$$\int_a^b \sqrt{f'^2(t) + g'^2(t)}\,dt. \tag{7}$$

Lebesgue observă că abia cu introducerea acestei definiții analitice a lungimii unei curbe, împreună cu definiția corespunzătoare pentru aria unei suprafețe, s-a desăvîrșit acea operă care este, de obicei, pusă în întregime pe seama lui Descartes, opera de reducere a geometriei la algebră (algebra într-un sens larg, vechi, al cuvîntului, cuprinzînd deci și analiza).

Definiția (7) pentru lungimea unei curbe, datorită lui Cauchy, definiție analitică, se referă la o clasă restrînsă de curbe. Dacă ținem seamă că pe vremea lui Cauchy singurele funcții integrabile erau cele continue (eventual cu un număr finit de puncte de discontinuitate), rezultă că, după această definiție, numai curbele pentru care funcțiile f și g au derivată continuă sînt înzestrate cu o lungime. Astfel de funcții f și g se numesc netede, iar despre curba corespunzătoare putem spune că este o curbă netedă.

Să facem cîteva observații — care ne vor fi utile — pe marginea definiției lui Cauchy.

Fiind dată o curbă netedă, parametrii directori ai tangentei la curbă sînt $f'(t)$ și $g'(t)$. Din expresia (7) se observă că lungimea unei curbe netede este o funcție continuă de $f'(t)$ și $g'(t)$. Iată ce înțelegem prin aceasta: dacă ne dăm un $\varepsilon > 0$, atunci există $\eta(\varepsilon) > 0$, astfel încît pentru orice curbă netedă $x = \varphi(t)$, $y = \psi(t)$, pentru care

$$|f'(t) - \varphi'(t)| < \eta(\varepsilon), \quad |g'(t) - \psi'(t)| < \eta(\varepsilon), \quad (a \leqslant t \leqslant b), \tag{8}$$

avem

$$\left| \int_a^b \sqrt{f'^2(t) + g'^2(t)}\,dt - \int_a^b \sqrt{\varphi'^2(t) + \psi'^2(t)}\,dt \right| < \varepsilon. \tag{9}$$

În adevăr, din inegalitatea $\sqrt{a^2+b^2} - \sqrt{c^2+d^2} \leqslant \sqrt{(a-c)^2+(b-d)^2}$ rezultă:

$$\left| \int_a^b (\sqrt{f'^2(t)+g'^2(t)} - \sqrt{\varphi'^2(t)+\psi'^2(t)})\, dt \right| \leqslant$$

$$\leqslant \left| \int_a^b \sqrt{[f'(t)-\varphi'(t)]^2+[g'(t)-\psi'(t)]^2}\, dt \right| .$$

Deci, de îndată ce inegalităţile (8) au loc pentru $\eta(\varepsilon) = \sqrt{\frac{\varepsilon}{2(b-a)}}$, se realizează inegalitatea (9). Aceasta înseamnă că lungimea unei curbe cu tangentă este o funcţie continuă de direcţia tangentei la curbă. Lungimea unei curbe cu tangentă încearcă o modificare oricît de mică dorim, cu condiţia ca tangenta la curbă să nu-şi modifice prea mult direcţia.

Să revenim acum la exemplul ilustrat în fig 7.III. În acest exemplu nu este îndeplinită condiţia de mai sus. În adevăr, fie un punct M situat pe segmentul BC, astfel încît perpendiculara ridicată în M pe BC să întîlnească orice linie poligonală din şir într-un punct în care aceasta admite tangentă unică. Deoarece însă tangenta la o astfel de linie poligonală are mereu înclinarea $\frac{\pi}{3}$ sau $\frac{2\pi}{3}$ radiani, cum, pe de altă parte, tangenta la BC are în orice punct înclinarea zero, rezultă că liniile poligonale construite nu tind în direcţie, ci doar în poziţie, către segmentul de dreaptă BC[1].

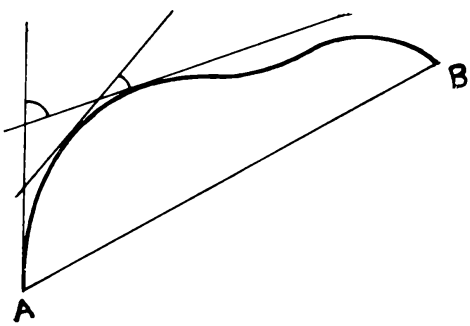


Fig. 9, III.

[1] A trebuit să luăm o anumită precauţie în alegerea lui M, deoarece liniile poligonale nu sînt curbe netede. Ele se descompun într-un număr finit de curbe netede. Însă precauţia poate fi înlăturată dacă lărgim puţin semnificaţia cuvintelor „tinde în direcţie" şi observăm că pentru un punct aparţinînd unei linii poligonale din şir, laturile (sau latura) liniei poligonale care conţin acel punct au, în mod constant, înclinările $\frac{\pi}{3}$ sau $\frac{2\pi}{3}$ radiani.

Fie acum Γ o curbă netedă. Fie P o linie poligonală înscrisă în Γ. O latură AB a lui P face cu tangentele la Γ, în punctele arcului AB al lui Γ, un unghi inferior celui mai mare unghi pe care-l fac între ele tangentele la Γ în punctele arcului AB (este suficient, pentru a ne da seama de aceasta, să facem uz de interpretarea geometrică a teoremei creșterilor finite; a se vedea fig. 9.III).

De aici rezultă că un șir P_n de linii poligonale înscrise în curba Γ, astfel încît lungimea celei mai mari laturi a lui P_n are ca limită pe zero, cînd $n \to \infty$, tinde către Γ nu numai în poziție (adică în sensul distanței introduse mai sus) ci și în direcție. Aceasta înseamnă că în corespondența care se stabilește între Γ și P_n nu numai că distanța între două puncte corespondente $A \in \Gamma$, $A_n \in P_n$ este inferioară unui $\varepsilon_n > 0$ ($\lim_{n\to\infty} \varepsilon_n = 0$), dar coeficientul unghiular al laturii lui P_n (sau al fiecăreia dintre laturile lui P_n), care conține pe A_n, diferă și el de coeficientul unghiular al tangentei la Γ în punctul A printr-o cantitate inferioară unui η_n, care tinde la zero cînd $n \to \infty$ [1].

În cazul circumferinței de pildă, luînd în rolul lui P_n șirul obținut de poligoane regulate, punctul A_n va fi situat la intersecția razei corespunzătoare punctului A de pe circumferință cu linia poligonală P_n.

În general: dacă un șir de linii poligonale înscrise într-o curbă cu tangentă tinde în poziție către această curbă, atunci el tinde și în direcție către aceeași curbă. Aceasta explică succesul teoriei lui Jordan și, în particular, corectitudinea rezultatului în teoria elementară a lungimii cercului, teorie care, după cum am văzut, abundă în confuzii și inadvertențe.

Șirurile de linii poligonale care tind către o curbă, atît în poziție, cît și în direcție, realizează, implicit, dezideratul (formulat de Lebesgue) asupra evitării drumurilor inutil complicate. Această situație oglindește faptul experimental constatat în orice măsurătoare: linia poli-

---

[1] Este de la sine înțeles că pentru ca toate aceste considerații să aibă sens și să nu comporte complicații, vom exclude, cel puțin în această discuție, curbele Γ care admit puncte singulare (în care $f'(t) = g'(t) = 0$). Considerațiile pe care le-am făcut se referă la curbe cu tangentă determinată.

gonală care înlocuiește curba, pentru cel care o măsoară, respectă, pe cît permite vederea omenească, direcția tangentei la curbă.

De ce atunci nu se definește lungimea unei curbe cu ajutorul liniilor poligonale ce o aproximează atît în poziție, cît și în direcție? Pentru curbe cu tangentă, acest mod de definire este nu numai corect, dar și cel mai indicat. Trebuie numai să demonstrăm că orice șir de linii poligonale care tinde către curbă, atît în poziție, cît și în direcție, furnizează aceeași limită pentru lungime. Aceasta se poate face fără mari complicații. Pentru curbe arbitrare însă, această definiție nu are sens. O curbă arbitrară, chiar simplă, poate să nu aibă tangentă în nici un punct. (Se știe, de pildă, că există funcții continue care nu sînt derivabile în nici un punct; graficul unei astfel de funcții este, deci, complet lipsit de tangentă).

Din cele văzute pînă acum rezultă că, în general, fiind dat un șir de curbe $\{\Gamma_n\}$ care tinde către o curbă Γ, șirul $\{l(\Gamma_n)\}$ nu tinde către $l(\Gamma)$. Este lesne de înțeles că șirurile $\{\Gamma_n\}$ de curbe rectificabile pentru care $l(\Gamma_n) \to l(\Gamma)$, cînd $n \to \infty$, Γ fiind curba limită, prezintă un deosebit interes. Aceste șiruri au condus la o noțiune nouă de convergență, numită convergența în lungime. Fie $\{f_n\}$ un șir de funcții continue și cu variație mărginită pe $[a, b]$, uniform convergent pe $[a, b]$. Șirul de curbe rectificabile

$$\Gamma_n : \begin{cases} x = t, \\ y = f_n(t), \end{cases} \qquad (a \leqslant t \leqslant b)$$

tinde către curba Γ, dată de $x = t$, $y = f(t)$, f fiind limita șirului f_n. Să presupunem că f este cu variație mărginită pe $[a, b]$.

Spunem că șirul de funcții $\{f_n\}$ converge în lungime către funcția f pe segmentul $[a, b]$, dacă $l[\Gamma_n]$ tinde către $l(\Gamma)$, atunci cînd $n \to \infty$. Se poate arăta că o condiție suficientă ca $\{f_n\}$ să conveargă în lungime către f este ca variația totală a funcției $f_n - f$ pe $[a, b]$ să tindă la zero, cînd $n \to \infty$.

Mai general, fiind date reprezentările parametrice

$$R_n : \begin{cases} x = f_n(t), \\ y = g_n(t). \end{cases} \qquad (a \leqslant t \leqslant b), \quad (n = 0, 1, 2, \ldots),$$

astfel încît șirul $f_1, f_2, f_3, \ldots$ tinde uniform pe $[a, b]$ către f_0, șirul $g_1, g_2, g_3, \ldots$ tinde uniform pe $[a, b]$ către g_0, iar f_n și g_n sînt funcții continue și cu variație mărginită pe $[a, b]$ pentru $n = 0, 1, 2, \ldots$, observăm că șirul de curbe $\{\Gamma_n\}$ definite de $\{R_n\}$ tinde către curba Γ_0 definită de R_0. Vom spune despre acest șir de reprezentări parametrice că converge în lungime pe $[a, b]$, dacă $\lim\limits_{n\to\infty} l(\Gamma_n) = l(\Gamma_0)$.

Se poate arăta că dacă șirul R_n converge în lungime către R_0, atunci șirul f_n converge în variație către f_0, iar șirul g_n converge în variație către g_0, pe $[a, b]$, înțelegînd prin aceasta că

$$\lim_{n\to\infty} \overset{b}{\underset{a}{V}} (f_n) = \overset{b}{\underset{a}{V}} (f_0), \qquad \lim_{n\to\infty} \overset{b}{\underset{a}{V}} (g_n) = \overset{b}{\underset{a}{V}} (g_0).$$

REPREZENTAREA INTEGRALĂ A LUNGIMII UNEI CURBE

Definiția dată de Cauchy lungimii unei curbe cu tangentă variind continuu merită, cu tot domeniul ei restrîns de valabilitate, să mai facă obiectul atenției noastre. Pentru a-i înțelege mai bine semnificația, să arătăm mai întîi că definiția lui Jordan conduce — pentru curbe netede — la expresia integrală a lui Cauchy. Fie deci curba Γ, dată de (1), cu $f'(t)$ și $g'(t)$ continue pe $[a, b]$.

Fie π o linie poligonală înscrisă în Γ, indusă de o diviziune

$$\Delta = (a = t_0 < \ldots < t_i < t_{i+1} < \ldots < t_n = b).$$

Lungimea lui π este dată de

$$l(\pi) = \sum_{i=0}^{n-1} \sqrt{[f(t_{i+1}) - f(t_i)]^2 + [g(t_{i+1}) - g(t_i)]^2}$$

sau încă, în virtutea teoremei creșterilor finite,

$$l(\pi) = \sum_{i=0}^{n-1} (t_{i+1} - t_i) \sqrt{f'^2(\xi_i) + g'^2(\eta_i)}$$
$$(t_i \leqslant \xi_i \leqslant t_{i+1}), \ (t_i \leqslant \eta_i \leqslant t_{i+1}).$$

Este ușor de verificat că

$$|\sqrt{f'^2(\xi_i)+g'^2(\eta_i)}-\sqrt{f'^2(t_i)+g'^2(t_i)}|\leqslant|f'(\xi_i)-f'(t_i)|+$$
$$+|g'(\eta_i)-g'(t_i)|,$$

deci

$$\left|l(\pi)-\sum_{i=0}^{n-1}(t_{i+1}-t_i)\sqrt{f'^2(t_i)+g'^2(t_i)}\right|\leqslant \tag{10}$$
$$\leqslant\sum_{i=0}^{n-1}(M_i-m_i)(t_{i+1}-t_i)+\sum_{i=0}^{n-1}(M_i'-m_i')(t_{i+1}-t_i),$$

unde M_i, m_i, M_i', m_i' sînt marginile funcțiilor f' și g' pe $[t_i, t_{i+1}]$.

În virtutea continuității uniforme a derivatelor f' și g' pe $[a, b]$, membrul din dreapta este inferior unui $\varepsilon>0$ dat, cu condiția ca norma diviziunii Δ să fie suficient de mică. Deci membrul din stînga din (10) tinde la zero pentru orice șir de diviziuni Δ_n de normă tinzînd la zero. De aici rezultă

$$l(\Gamma)=\int_a^b\sqrt{f'^2(t)+g'^2(t)}\,dt. \tag{11}$$

Am desfășurat, în mod intenționat, toate detaliile demonstrației. O privire superficială asupra acestei demonstrații lasă impresia că intervenția continuității derivatelor este esențială. O analiză atentă pune însă în evidență un fapt care reabilitează aproape în întregime definiția lui Cauchy: expresia integrală a lungimii este valabilă pentru o clasă mult mai largă de curbe rectificabile decît clasa curbelor netede, cu condiția să lucrăm cu o noțiune mai generală de integrabilitate.

Iată acum etapele prin care a ajuns știința la acest rezultat. Mai întîi, Riemann a generalizat noțiunea de integrabilitate, efectuînd următoarea construcție: fie pentru orice diviziune $\Delta=(a=t_0<\ldots<t_i<t_{i+1}<\ldots<t_n=b)$ și pentru f, definită pe $[a, b]$, expresia

$$\sigma(\Delta)=\sum_{i=0}^{n-1}f(\xi_i)(t_{i+1}-t_i),\qquad(t_i\leqslant\xi_i\leqslant t_{i+1}).$$

Fie $\nu(\Delta)$ norma lui Δ. Dacă există un număr I, astfel încît pentru orice $\varepsilon > 0$ se poate determina un $\eta > 0$, cu proprietatea că de îndată ce $\nu(\Delta) < \eta$, avem

$$|\sigma(\Delta) - I| < \varepsilon,$$

oricare ar fi alegerea valorilor ξ_i în $[t_i, t_{i+1}]$, atunci f este integrabilă în sens Riemann pe $[a, b]$, integrala lui f pe $[a, b]$ fiind I. Darboux a construit aceeași integrală pe o altă cale, și anume: fie o funcție f mărginită pe $[a, b]$. Fie m_i, M_i marginile lui f pe $[t_i, t_{i+1}]$. Fie

$$s(\Delta) = \sum_{i=0}^{n-1} m_i(t_{i+1} - t_i), \quad S(\Delta) = \sum_{i=0}^{n-1} M_i(t_{i+1} - t_i),$$

avem

$$s(\Delta) \leqslant S(\Delta).$$

Se poate arăta că o condiție necesară și suficientă ca f să fie integrabilă Riemann pe $[a, b]$ este ca, pentru fiecare $\varepsilon > 0$, să se poată determina o diviziune Δ a lui $[a, b]$, astfel încît $S(\Delta) - s(\Delta) < \varepsilon$. De aici rezultă, ținînd seama că în membrul al doilea din (10) se află tocmai diferența S—s, referitoare întîi la f', apoi la g', că relația (11) este valabilă ori de cîte ori derivatele f' și g' sînt integrabile Riemann. Această formulare poate să pară curioasă acelor cititori care nu sînt obișnuiți să lucreze cu funcții integrabile, altele decît cele continue. Pentru funcțiile continue, nu distingem între existența primitivei și existența integralei. O funcție continuă este în același timp integrabilă și derivată. Există însă derivate mărginite care nu sînt integrabile Riemann. Matematicianul român D. Pompeiu a definit o întreagă clasă de funcții a căror derivată este mărginită, dar neintegrabilă Riemann. Așadar, expresia „derivată integrabilă Riemann" nu este un pleonasm. Nu mai insistăm aici asupra acestor lucruri, deoarece integrala Riemann este reluată și discutată amănunțit în capitolul IV.

NECESITATEA DE A SE INTRODUCE O NOUĂ NOȚIUNE DE INTEGRALĂ: INTEGRALA LUI LEBESGUE

Dar lucrurile nu s-au oprit aici. Lebesgue a avut ideea de a controla dacă nu cumva derivabilitatea funcțiilor f și g nu rezultă din însăși structura unei curbe rectificabile. Lebesgue a demonstrat că dacă f și g sînt funcțiile care dau reprezentarea parametrică a unei curbe rectificabile, atunci f și g sînt derivabile în fiecare punct, exceptînd eventual o mulțime de măsură nulă de puncte. (O mulțime de puncte de pe dreaptă este de măsură nulă dacă pentru fiecare $\varepsilon > 0$ se poate găsi o familie cel mult numărabilă de intervale de lungime totală mai mică decît ε, și astfel încît orice punct al mulțimii să fie conținut într-un interval al familiei.) Tot Lebesgue a introdus o noțiune nouă de integrabilitate, care-i poartă numele, și care permite integrarea oricărei derivate dintr-o funcție care intră în reprezentarea parametrică a unei curbe rectificabile. Această noțiune de integrabilitate nici nu cere funcției care se integrează pe $[a, b]$ să fie definită pe tot $[a, b]$; ea poate să nu fie definită pe o mulțime de măsură nulă de puncte (așa cum se întîmplă de fapt, în general, cu derivatele funcțiilor f și g de mai sus).

Nu putem da aici definiția constructivă a integralei lui Lebesgue (ea va fi prezentată în capitolul IV). Vom da însă o definiție descriptivă a acestei integrale, definiție echivalentă cu cea constructivă. Dar mai întîi, o definiție pregătitoare.

O funcție φ definită pe $[a, b]$ este absolut continuă pe $[a, b]$, dacă fiecărui $\varepsilon > 0$ îi corespunde un $\eta > 0$, astfel încît dacă (a_1, b_1), (a_2, b_2) ... (a_n, b_n) formează un sistem finit de intervale disjuncte din $[a, b]$, cu $\sum_{i=1}^{n} (b_i - a_i) < \eta$, atunci $\sum_{i=1}^{n} f(b_i) - f(a_i) < \varepsilon$.

Se poate arăta că orice funcție absolut continuă pe $[a, b]$ este continuă și cu variație mărginită pe $[a, b]$. Reciproca, însă, nu este adevărată.

Orice funcție cu derivată mărginită pe $[a, b]$ este absolut continuă pe $[a, b]$.

Spunem că funcția f, definită pe $[a, b]$, este integrabila Lebesgue pe $[a, b]$, dacă există o funcție F absolut continuă pe $[a, b]$ și o mulțime E, de măsură nulă, din $[a, b]$, astfel încît $F'(t) = f(t)$, pentru orice t din $[a, b]$ care nu aparține mulțimii E. Numărul $F(b) - F(a)$ se numește integrala în sensul lui Lebesgue a funcției f de la a la b.

Se constată ușor că integrala Riemann, ca funcție de limita superioară de integrare, este absolut continuă. Pe de altă parte, în orice tratat de analiză se arată că

$$F(x) = \int_a^x f(t)\mathrm{d}t$$

are ca derivată, în orice punct x_0 în care $f(t)$ este continuă, chiar pe $f(x_0)$. Deoarece am semnalat mai sus că o funcție integrabilă Riemann este continuă în fiecare punct, cu excepția unei mulțimi de puncte de măsură nulă , rezultă că o funcție integrabilă Riemann este integrabilă Lebesgue, și valorile celor două integrale coincid.

S-ar putea pune problema dacă integrala Lebesgue a unei funcții este unic determinată de această funcție. Răspunsul este afirmativ, deoarece se poate demonstra că dacă două funcții absolut continue au în fiecare punct, exceptînd o mulțime de măsură nulă, aceeași derivată, atunci ele diferă printr-o constantă.

Iată acum formularea rezultatelor care încoronează considerațiile privitoare la reprezentarea integrală a lungimii unei curbe:

Pentru orice curbă rectificabilă (dată de (1)) derivatele funcțiilor f și g, ca și expresia $\sqrt{f'^2(t) + g'^2(t)}$, sînt integrabile în sensul lui Lebesgue pe $[a, b]$. În general, avem

$$l(\Gamma) \geqslant \int_a^b \sqrt{f'^2(t) + g'^2(t)}\, \mathrm{d}t, \qquad (12)$$

integrala fiind luată în sensul lui Lebesgue. Necesar și suficient ca în (12) să apară semnul $=$ este ca f și g să fie funcții absolut continue pe $[a, b]$.

Acum se explică din ce cauză pentru curbe date de funcții f și g, cu derivate integrabile Riemann, avem relația (11). Astfel de funcții sînt absolut continue, deci intră exact în cazul în care relația (12) devine o egalitate.

LUNGIMEA UNEI CURBE, CA FUNCȚIONALĂ INFERIOR SEMICONTINUĂ ÎN SPAȚIUL CURBELOR. PUNCTUL DE VEDERE AL LUI FRÉCHET

Se spune uneori că definițiile sînt libere. Noi n-am adoptat aici această concepție și am căutat să justificăm alegerea fiecărei definiții. Vom prezenta acum un alt aspect, pus în evidență de Maurice Fréchet, al noțiunii de lungime a unei curbe, așa cum a fost ea definită de Lebesgue. Prezentarea lui Fréchet înlătură orice element arbitrar care ar mai putea fi atribuit definiției date de Lebesgue.

Să considerăm mulțimea C a tuturor curbelor definite de (1) cu f și g continue pe $[a, b]$. Vom numi C spațiul curbelor. Datorită definițiilor introduse pînă aici, știm ce trebuie să înțelegem prin afirmația: „Un șir de curbe $\{\Gamma_n\}$ tinde în poziție către curba Γ".

Să presupunem acum că definim pe spațiul C o funcție reală φ. Deci fiecărei curbe $\Gamma \in C$ i se asociază, în mod univoc, un număr real $\varphi(\Gamma)$. (Admitem și posibilitatea ca $\varphi(\Gamma)$ să fie infinit.) Funcția φ se numește o funcțională definită pe spațiul C. Curbele Γ vor fi numite „puncte" ale spațiului C.

Vom spune că funcționala φ este continuă în punctul Γ, dacă $\varphi(\Gamma_n) \to \varphi(\Gamma)$ de îndată ce Γ_n tinde în poziție către Γ. Să presupunem că φ este funcționala definită de lungimea curbei Γ. φ este, în acest caz, o funcțională nenegativă definită pe C, dar din cele arătate în acest capitol rezultă că această funcțională nu este continuă în nici un punct din C.

Pentru a putea caracteriza lungimea lui Γ, ca funcțională definită pe C, vom recurge la o noțiune mai slabă decît continuitatea, anume la noțiunea de semicontinuitate. Funcționala φ, definită pe C, este inferior (respectiv, superior) semicontinuă în punctul $\Gamma \in C$, dacă pentru orice șir $\Gamma_n \to \Gamma$, pentru care există $\lim_{n\to\infty} \varphi(\Gamma_n)$, avem

$$\lim_{n\to\infty} \varphi(\Gamma_n) \geqslant \varphi(\Gamma) \text{ (respectiv } \lim_{n\to\infty} \varphi(\Gamma_n) \leqslant \varphi(\Gamma)). \quad (13)$$

Este vizibil că o funcțională inferior și superior semicontinuă în punctul Γ este pur și simplu continuă în punctul Γ și reciproc [1].

Forma (13) a definiției semicontinuității este cea mai convenabilă în problema care ne interesează, pentru că ea pune în evidență următorul fapt, deosebit de important: funcționala care definește lungimea unei curbe (a se vedea varianta Lebesgue) este o funcțională inferior semicontinuă în fiecare punct din spațiul C. (Putem considera că această funcțională este definită pe tot spațiul C, deoarece orice curbă are lungime, finită sau infinită, și pe de altă parte, am acceptat pentru φ și valori infinite.)

În spațiul C există unele „puncte" care atrag în mod deosebit atenția: sînt punctele Γ care reprezintă un caz particular de curbe, anume liniile poligonale. Să notăm cu P mulțimea acestor puncte din C. Avem $P \subset C$. Funcționala φ prezintă o situație specială pe P, prin faptul că pe P, φ coincide cu lungimea definită pe cale elementară, ca suma lungimilor segmentelor care compun linia poligonală respectivă. Putem spune deci:

Lungimea unei curbe este o funcțională nenegativă, inferior semicontinuă pe C și care se reduce pe $P \subset C$ la lungimea în sens elementar.

Problema definirii lungimii unei curbe apare, astfel, ca o problemă de extensiune, de prelungire a unei funcționale definite pe P, la o funcțională definită pe C, prelungire care să se facă cu respectarea anumitor cerințe intuitive fundamentale, legate de noțiunea de lungime. Meritul lui Lebesgue este de a fi identificat o astfel de cerință și de a o fi reflectat în proprietatea corespunzătoare semicontinuității inferioare, proprietate care constituie însăși esența definiției pe care el a pus-o.

Se poate demonstra că lungimea liniilor poligonale, definită elementar ca sumă a lungimilor segmentelor care compun o astfel de linie, este o funcțională inferior semicontinuă pe P. În felul acesta, semicontinuitatea

---

[1] De obicei, definiția semicontinuității nu se formulează ca mai sus, ci într-o altă variantă, echivalentă cu cea dată aici.

inferioară a lungimii pe C, semicontinuitate care rezultă fără dificultăți din definiția lui Lebesgue, prelungește o proprietate a lungimii liniilor poligonale.

Acest procedeu de „prelungire" nu este întîlnit aici pentru prima dată. Pentru definirea măsurii unui segment se definește mai întîi măsura segmentelor comensurabile cu unitatea de lungime, apoi se impune măsurii segmentelor oarecare condiția de a fi o funcție continuă, egală, pentru segmente comensurabile, cu măsura deja definită. Schema comună tuturor acestor cazuri este: Se prelungește o funcție de la un cîmp dat la altul, mai vast, prin construcția unei funcții egale cu funcția dată pe cîmpul dat și păstrînd pe cîmpul mai întins unele proprietăți, pe care le socotim mai importante, ale funcției date.

În legătură cu ultima formă pe care am dat-o definiției lungimii unei curbe, se ridică următoarea problemă: Este adevărată afirmația reciprocă, anume că orice funcțională nenegativă, inferior semicontinuă pe C, care se reduce pe $P \subset C$ la lungimea în sens elementar, este însăși funcționala care furnizează lungimea curbelor $\Gamma \in C$?

Fréchet a observat că răspunsul la această întrebare este negativ și și-a pus problema de a caracteriza, în mulțimea funcționalelor cu proprietățile arătate, definite pe C, pe aceea care furnizează lungimea. El a reușit să rezolve această problemă, demonstrînd că dintre toate funcționalele φ inferior semicontinue pe C, reducîndu-se la lungimea obișnuită pe P, funcționala l, care dă lungimea curbelor, este cea mai puțin discontinuă, în sensul că are în fiecare punct oscilația „cea mai mică". Deoarece limita superioară a oricărei funcționale φ cu proprietățile arătate este, în fiecare punct din C, egală cu ∞, condiția găsită de Fréchet revine la inegalitatea $l(\Gamma) \geqslant \varphi(\Gamma)$ pentru orice φ și orice $\Gamma \in C$. Pentru a o demonstra, este suficient să observăm că există un șir (π_n) de puncte din P, astfel încît $\lim_{n\to\infty} \pi_n = \Gamma$ și $l(\pi_n) \to l(\Gamma)$. Pe de altă parte, avem $\varphi(\pi_n) = l(\pi_n)$, deci limita inferioară a funcției φ în punctul Γ este mai mică sau egală cu $l(\Gamma)$ și, ca urmare a semicontinuității inferioare, $\varphi(\Gamma) \leqslant l(\Gamma)$.

Să dăm acum un exemplu de funcție φ inferior semicontinuă pe C, care se reduce pe P la lungimea obișnuită și care nu coincide pe C cu lungimea Lebesgue.

Să considerăm mulțimea $S \subset C$ formată din toate semicircumferințele de rază egală cu unitatea. Funcția φ, egală pe $C - S$ cu lungimea Lebesgue și egală pe S cu jumătate din lungimea Lebesgue, îndeplinește condițiile dorite.

IV

Itinerar în teoria integralei

INTRODUCERE

Una dintre problemele care i-a frămîntat pe matematicieni din cele mai vechi timpuri este aceea a determinării ariei unei porțiuni din plan. Forma cea mai simplă a acestei probleme este aceea a determinării ariei pătratului de latură 1. Se convine a se considera că aria unui astfel de pătrat este egală cu 1. Sensul acestei afirmații este următorul: dacă latura unui pătrat este de 1 cm, atunci aria lui este de 1 cm^2. Cu alte cuvinte, ne alegem drept unitate de măsură pentru arii aria unui pătrat a cărui latură este unitatea de lungime. O problemă care se rezolvă pe o cale destul de elementară pentru a o putea aborda încă în școala elementară este aceea a determinării ariei unui dreptunghi ale cărui laturi se exprimă prin numere raționale. Urmează apoi o problemă care prezintă un salt calitativ față de problema precedentă: aceea a determinării ariei unui pătrat a cărui latură nu se mai exprimă printr-un număr rațional. Aici, metodele matematicii elementare se dovedesc neputincioase. Pentru a determina aria unui astfel de pătrat este nevoie de o operație nouă, specifică analizei matematice: operația de trecere la limită. Într-adevăr, dacă latura unui pătrat se exprimă printr-un număr a incomensurabil (adică irațional), se poate determina, prin însăși definiția acestor numere, un șir de numere raționale a_n astfel încît $\lim a_n = a$. Aria pătratului de latură a va fi tocmai limita șirului $a_1^2, a_2^2, \ldots, a_n^2, \ldots$ După cum vedem, posibilitatea de a determina aria oricărui pătrat este intim legată de cunoașterea noțiunii de număr real, în toată generalitatea

ei. În mod asemănător se determină și aria unui dreptunghi pentru care cel puțin una din laturi se exprimă printr-un număr irațional. Cunoscînd aria unui dreptunghi, putem determina imediat aria oricărui triunghi, deci și aria oricărui poligon, deoarece un poligon se descompune într-un număr finit de triunghiuri.

Am trecut în fugă peste toate aceste lucruri, pentru că ele ne sînt binecunoscute încă din școala medie, unele dintre ele chiar din școala elementară.

Acest mod de a studia aria unui poligon urmează oarecum calea istorică, adică reproduce, într-o formă sintetică, etapele istorice prin care a trecut cunoașterea noțiunii de arie pentru poligoane. Însă, atunci cînd oamenii au acumulat o experiență destul de îndelungată în problema determinării ariei, toate considerațiile acceptate de foarte multă vreme și intrate oarecum în obișnuința de fiecare zi a matematicienilor au început să fie supuse unui riguros examen critic. S-a pus problema solidității, din punct de vedere logic, a cunoștințelor noastre despre aria unui poligon.

În teoria clasică a ariei, noțiunea de arie a unui poligon este considerată cunoscută și se pune doar problema determinării numerice a ariei. Mai mult decît atît, sînt admise o seamă de proprietăți fundamentale ale ariei, de exemplu: pătratul de latură 1 are aria egală cu 1; două poligoane egale au aceeași arie; dacă poligonul P este reuniunea a două poligoane p' și p'', fără puncte interioare comune, atunci aria lui P este egală cu suma dintre aria lui p' și aria lui p''.

Desigur că în postularea existenței ariei pentru orice poligon, cît și în postularea proprietăților de mai sus ale ariei nu trebuie să vedem un act arbitrar al unor matematicieni. Astfel de postulate nu fac decît să sintetizeze niște observații pe care oamenii le-au efectuat de miliarde și miliarde de ori. Aceste postulate sînt foarte prețioase, pentru că prin ele se întreține, în bună măsură, contactul matematicii cu natura, cu realitatea vie, cu experiența și practica milenară a oamenilor. Se poate totuși pune problema dacă postulatele au fost bine alese, cu alte cuvinte dacă ele exprimă corect și complet conținutul real, intuitiv,

pe care sînt chemate să-l oglindească. Din acest punct de vedere, s-au mai adus uneori — și poate se vor mai aduce și în viitor — unele modificări, de detaliu, postulatelor de mai sus.

CUM DETERMINĂM ARIA UNEI FIGURI MAI COMPLICATE?

Să încercăm să determinăm aria unei porțiuni din plan mai complicată decît un poligon. Aceasta înseamnă că porțiunea considerată nu este delimitată numai de segmente de dreaptă, ci și de arce de curbă. Această problemă nu este chiar așa de nouă; am mai întîlnit-o atunci cînd am determinat aria unui cerc. Totuși, dacă cineva ne-ar cere să transpunem metoda de acolo în alte cazuri, de pildă să determinăm cu ajutorul ei aria unei elipse, ne-am găsi într-o încurcătură foarte mare. Oricît de bine am cunoaște procedeul de determinare a ariei unui cerc, încercarea de a transpune acest procedeu la elipsă nu va reuși, deoarece el utilizează proprietăți caracteristice ale cercului, proprietăți pe care elipsa nu le are. Într-adevăr, elipsa nu mai poate fi aproximată oricît de bine cu poligoane regulate.

Dificultatea pe care am întîlnit-o mai sus este caracteristică pentru o întreagă perioadă, foarte lungă, de dezvoltare a teoriei ariei. Un timp foarte îndelungat, matematicienii știau să determine, prin diverse mijloace, aria anumitor figuri particulare, dar nu reușeau să elaboreze o metodă generală, un procedeu uniform de determinare a ariei oricărei porțiuni din plan, adică un procedeu care să nu angajeze particularitățile figurii. Un exemplu celebru în acest sens este modul în care marele matematician al Greciei antice, Arhimede, determina aria unui segment de parabolă. Să considerăm arcul de parabolă ACB (vezi fig. 1.IV) de ecuație $y^2 = x$.

Fie $AA'B'BA$ dreptunghiul construit, așa cum se vede în figură. Fie a_1 aria triunghiului ABC. Fie D' și E' mijloacele laturilor AC și BC. Fie D și E punctele de intersecție cu parabola ale paralelelor prin D' și E' la CC'. Fie a_2 suma ariilor triunghiurilor ACD și BCE.

În cele patru segmente parabolice rămase construim, după același procedeu, patru triunghiuri și notăm cu a_3 suma ariilor acestor patru triunghiuri. Continuînd astfel, observăm că suma $a_1 + a_2 + a_3 + \ldots + a_n$ se apropie

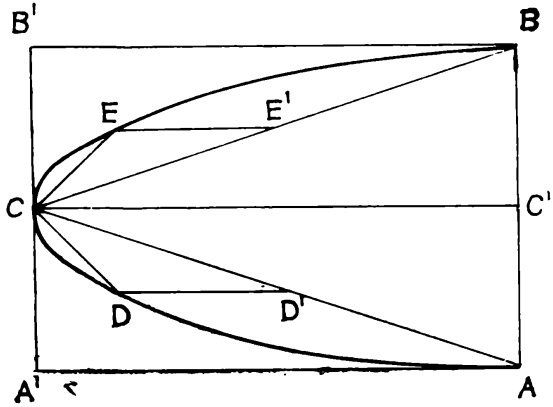


Fig. 1.IV

cu atît mai mult de aria a a segmentului parabolic (cuprins între arcul ACB și coarda AB) cu cît n este mai mare. Avem deci:

$$a = \lim_{n\to\infty} (a_1 + a_2 + a_3 + \ldots + a_n) = a_1 + a_2 + a_3 + \ldots + a_n + \ldots .$$

Însă avem $DD' = EE' = \frac{CC'}{4}$, deoarece DD' este diametrul conjugat direcției AC, iar EE' este diametrul conjugat direcției BC. Pe de altă parte, raportul dintre aria triunghiului ACD și aria triunghiului ACC' este egal cu raportul înălțimilor. Deducem deci:

$$a_2 = \frac{a_1}{4}, \quad a_3 = \frac{a_2}{4}, \ldots, a_{n+1} = \frac{a_n}{4}, \ldots,$$

de unde

$$a = a_1 + \frac{a_1}{4} + \frac{a_1}{4^2} + \frac{a_1}{4^3} + \ldots + \frac{a_1}{4^n} + \ldots = \frac{a_1}{1 - \frac{1}{4}} = \frac{4a_1}{3},$$

deci aria segmentului parabolic cuprins între arcul $ADCEB$ și coarda AB este egală cu două treimi din aria dreptunghiului $AA'B'BA$.

Ingeniozitatea procedeului de mai sus nu poate fi contestată. Totuși, caracterul foarte special, atît de legat de particularitățile parabolei, face ca el să nu poată fi extins și la alte cazuri.

DE LA ARHIMEDE LA CAVALIERI

Este interesant că Arhimede a găsit încă două procedee de determinare a ariei unui segment de parabolă. Unul dintre ele constă în a reduce problema la aceea a determinării centrului de greutate al unui triunghi, segmentul de parabolă fiind asimilat cu o sumă infinită de segmente de dreaptă paralele cu axa.

Tot Arhimede a determinat aria porțiunii limitată de spirala care-i poartă numele (a cărei ecuație este $\rho = C\omega$), iar înaintea sa Eudox determinase volumul piramidei și al conului, ale căror formule fuseseră date, fără demonstrație, de Democrit. Antichitatea greacă a cunoscut de asemenea modul de determinare a lungimii cercului, a ariei unui segment sferic, ca și a momentelor statice. Toate aceste probleme erau puse de geometrie și de statică, științe relativ mai dezvoltate în această epocă. Aceste determinări făcute de cei mai strălucițti matematicieni ai Greciei antice conțin, fără îndoială, germenele a ceea ce avea să se numească mai tîrziu „calculul integral". În nici un caz însă nu putem vorbi de un calcul integral în antichitate. O trăsătură caracteristică a calculului integral este sesizarea faptului că o seamă de determinări de arii, volume, momente, în ciuda diversității pe care o prezintă ele din punctul de vedere al naturii mărimilor respective, conduc la una și aceeași problemă matematică, la una și aceeași schemă. De exemplu, problema determinării volumului piramidei, rezolvată de Eudox, a determinării ariei unui segment de parabolă, a centrului de greutate a unui triunghi și a ariei unei spirale, rezolvate de Arhimede, se reduc, toate la determinarea limitei unei sume de forma $\xi_1^2 +$

$+ \xi_2^2 + \ldots + \xi_n^2$ atunci cînd n tinde la infinit, cu alte cuvinte, în notația și limbajul de astăzi, toate cele patru probleme de mai sus se reduc la calculul unei integrale de forma

$$\int_a^b x^2 \, dx.$$

Or, tocmai conștiința acestei analogii le lipsea matematicienilor din antichitate. Separarea formei de conținut s-a dovedit a fi un proces istoric foarte lent, în care progresele au venit încetul cu încetul.

După lunga perioadă de amorțire, de stagnare a științei, caracteristică evului mediu, procesul de cristalizare a noțiunii de integrală avea să cunoască un progres esențial abia la sfîrșitul secolului al XVI-lea și, mai ales, în secolul al XVII-lea. În această perioadă, se stabilesc procedee de determinare a ariei unei suprafețe de revoluție și, sub impulsul lucrărilor lui Huygens asupra pendulului compus, se calculează pentru prima oară momente de inerție. Acum situația era coaptă pentru ca matematicienii să înceapă să vadă legătura între toate aceste probleme de geometrie și de mecanică. Pentru arii și volume, această analogie este pusă în evidență cu o deosebită pregnanță, în prima jumătate a secolului al XVII-lea, de către Cavalieri în a sa *Geometrie a indivizibilelor*. Cavalieri enunță și încearcă să stabilească principiul următor: „Dacă două arii plane sînt astfel încît orice paralelă la o direcție dată le taie după segmente ale căror lungimi sînt într-un raport constant, atunci aceste arii sînt în același raport". Un principiu analog este enunțat pentru „volumele tăiate de planele paralele cu un plan fix după arii ale căror măsuri sînt într-un raport constant". În absența unui limbaj matematic adecvat, Cavalieri exprimă aici într-un mod necorespunzător o observație corectă: „Fie două arii, una definită de inegalitățile

$$0 \leqslant x \leqslant a, \; 0 \leqslant y \leqslant f(x),$$

iar cealaltă definită de inegalitățile

$$0 \leqslant x \leqslant a, \; 0 \leqslant y \leqslant g(x).$$

Sumele de ordonate

$$f(0) + f\left(\frac{a}{n}\right) + f\left(\frac{2a}{n}\right) + \cdots + f\left(\frac{(n-1)a}{n}\right)$$

$$g(0) + g\left(\frac{a}{n}\right) + g\left(\frac{2a}{n}\right) + \cdots + g\left(\frac{(n-1)a}{n}\right),$$

sînt, una față de alta, într-un raport care, pentru n destul de mare, se apropie oricît de mult de raportul celor două arii". Cavalieri trece la limită pentru $n \to \infty$ și spune că suma tuturor ordonatelor primei curbe este, față de suma tuturor ordonatelor celei de a doua curbe, într-un raport riguros egal cu raportul ariilor.

UN ALT MOD DE A DETERMINA ARIA UNUI SEGMENT DE PARABOLĂ

Să ne întoarcem din nou la problema determinării ariei unui segment de parabolă și să încercăm s-o rezolvăm pe o cale care să angajeze cît mai puțin particularitățile acestei curbe. Fie deci parabola de ecuație $y = x^2$. Aria porți-

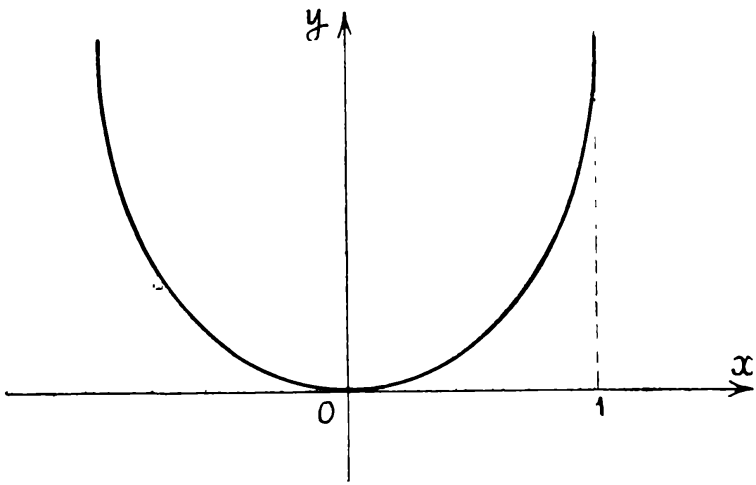


Fig. 2.IV

unii cuprinse în interiorul acestei parabole este, evident, infinită. Să limităm atunci această porțiune cu ajutorul unei drepte, de exemplu, dreapta de ecuație $x = 1$ (vezi fig 2.IV).

Însă, datorită simetriei parabolei, considerate în raport cu axa Ox, va fi suficient să calculăm aria porțiunii cuprinse între ramura superioară a parabolei axa Ox și dreapta $x = 1$. Dublînd rezultatul obținut, vom obține aria întregii porțiuni. În felul acesta, problema s-a redus la aceea a determi-

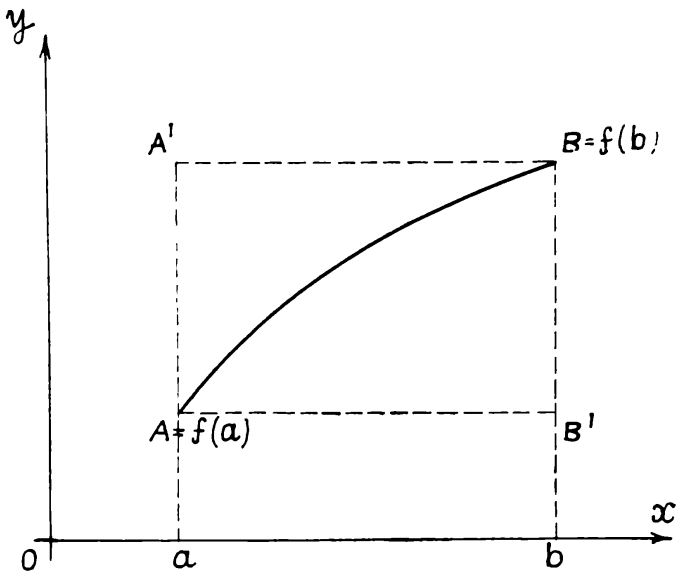


Fig. 3.IV

nării ariei porțiunii delimitate de graficul funcției $f(x) = x^2$, axa Ox și dreapta $x = 1$. Funcția $f(x) = x^2$ este crescătoare și pozitivă pe intervalul $[0, 1]$. Problema pe care trebuie s-o rezolvăm este deci un caz particular al următoarei probleme: fie $f(x)$ o funcție pozitivă și crescătoare pe intervalul $[a, b]$. Să determinăm aria α a porțiunii din plan cuprinse între graficul funcției $f(x)$, axa Ox și dreptele $x = a$ și $x = b$ (vezi fig. 3. IV).

Într-o primă aproximație, am putea lua drept arie a porțiunii considerate aria unuia dintre dreptunghiurile $AabB'$ și $A'Bba$ (vezi fig. 3. IV).

Eroarea pe care o comitem în această aproximare a ariei este inferioară diferenței ariilor celor două dreptunghiuri, diferența egală cu

$$(f(b) - f(a))(b - a).$$

Această majorare a erorii nu este însă satisfăcătoare, deoarece ea nu este la dispoziția noastră. Atît mărimea $b - a$, cît și $f(b) - f(a)$ depind de datele problemei, deci de aceste date ale problemei și nu de noi depinde faptul dacă eroarea comisă este, de pildă, mai mică sau nu decît 10^{-2}. Este firesc atunci să împărțim intervalul $[a, b]$ în intervale parțiale prin punctele

$$a = x_0 < x_1 < x_2 < \ldots < x_i < x_{i+1} < \ldots x_n = b,$$

și să considerăm drept valoare aproximativă a ariei α suma s a ariilor dreptunghiurilor avînd bazele egale cu $x_1 - x_0$, $x_2 - x_1$, ..., $x_{i+1} - x_i$, ..., $x_n - x_{n-1}$ și înălțimile egale, respectiv, cu $f(x_0)$, $f(x_1)$, ..., $f(x_i)$, ..., $f(x_{n-1})$.

Este ușor de văzut că

$$\alpha - s \leqslant (f(x_1) - f(a))\,(x_1 - a) + (f(x_2) - f(x_1))\,(x_2 - x_1) + \\ + \ldots + (f(x_{i+1}) - f(x_i))\;\;(x_{i+1} - x_i) + \ldots + \\ + (f(b) - f(x_{n-1}))\,(b - x_{n-1}).$$

Să presupunem acum că dorim ca eroarea $\alpha - s$ să nu depășească numărul 10^{-2}. Pentru a realiza aceasta, vom alege punctele $x_1, x_2, \ldots, x_{n-1}$ destul de apropiate pentru ca fiecare dintre diferențele

$$f(x_1) - f(a),\ f(x_2) - f(x_1),\ \ldots,\ f(x_{i+1}) - f(x_i),\ \ldots,$$
$$f(b) - f(x_{n-1}),$$

să fie inferioară lui $10^{-2}/(b - a)$. În acest fel obținem

$$\alpha - s \leqslant \frac{10^{-2}}{b-a}[(x_1 - a) + (x_2 - x_1) + \ldots + (x_{i-1} - x_i) + \\ + \ldots + (b - x_{n-1})] = \frac{10^{-2}}{b-a}(b - a) = 10^{-2}$$

și problema micșorării erorii este rezolvată. Totul depinde, după cum se vede, de alegerea potrivită a punctelor $x_1, x_2, \ldots, x_i, \ldots, x_{n-1}$.

Să aplicăm procedeul expus funcției $f(x) = x^2$, considerate pe intervalul $[0, 1]$. Pentru simplificare, să alegem punctele $x_1, x_2, \ldots, x_i, \ldots, x_{n-1}$, astfel încît

$$x_1 - a = x_2 - x_1 = \ldots = x_{i+1} - x_i = \ldots = b - x_{n-1} = \frac{1}{n}.$$

Avem

$$x_{i+1}^2 - x_i^2 = (x_{i+1} - x_i)(x_{i+1} + x_i) = \frac{x_{i+1} + x_i}{n} < \frac{2}{n} \quad (i = 0, 1, \ldots, n-1)$$

deci, pentru ca $x_{i+1}^2 - x_i^2 < 10^{-2}$, oricare ar fi $i = 0, 1, \ldots, \ldots, n-1$, este suficient ca $10^{-2} n > 2$, adică $n > 200$. Deci, de îndată ce $n > 200$, expresia

$$\frac{1}{n}\left(\frac{1}{n^2} + \frac{2^2}{n^2} + \frac{3^2}{n^2} + \cdots + \frac{i^2}{n^2} + \cdots + \frac{(n-1)^2}{n^2}\right)$$

aproximează aria porțiunii delimitate de arcul $y = x^2$, axa Ox și dreapta $x = 1$ cu o eroare mai mică decît 10^{-2}.

Este ușor de înțeles acum că în loc de 10^{-2} putem lua orice număr dorim, oricît de mic. Aria porțiunii considerate va putea fi aproximată cu o eroare mai mică decît acest număr, cu condiția ca n să fie destul de mare. Valoarea exactă a ariei căutate va fi

$$\lim_{n \to \infty} \frac{1}{n}\left[\frac{1}{n^2} + \frac{2^2}{n^2} + \frac{3^2}{n^2} + \cdots + \frac{i^2}{n^2} + \cdots + \frac{(n-1)^2}{n^2}\right] = \frac{1}{3}.$$

CAZUL ÎN CARE FUNCȚIA NU MAI ESTE MONOTONĂ

Ușurința cu care am determinat, în cele de mai sus, valoarea ariei cu o eroare oricît de mică o doream se datora faptului că arcul de curbă care delimita porțiunea de plan considerată era graficul unei funcții monotone. Ce se întîmplă însă dacă funcția nu este monotonă? La prima ve-

dere, s-ar părea că orice funcție continuă este alcătuită din „bucăți monotone". Dacă ar fi așa, ar rămîne doar să descompunem intervalul $[a, b]$ în intervale parțiale, astfel încît pe fiecare interval parțial funcția să fie monotonă și să aplicăm pe fiecare interval de monotonie considerațiile expuse mai sus. Însă analiștii au pus în evidență, încă în secolul trecut, existența funcțiilor continue care nu sînt monotone pe nici un interval. Deci, din procesul de determinare a ariei trebuie să eliminăm ipoteza de monotonie. Acest lucru nu va fi atît de greu, dacă ne gîndim că monotonia funcției a intervenit doar în chestiunea evaluării erorii. Vom menține deci procedeul utilizat la funcții monotone în chestiunea determinării aproximative a ariei, urmînd ca evaluarea erorii să constituie obiectul unei alte cercetări. De altfel, vom avea grijă să marcăm momentul în care aceste două probleme se despart, pornind fiecare pe drumul ei propriu.

Fie $f(x)$ o funcție pozitivă pe intervalul $[a, b]$. Fie D porțiunea limitată de graficul funcției $f(x)$, axa absciselor și paralele $x = a$ și $x = b$ la axa ordonatelor. Să împărțim intervalul $[a, b]$ în intervale parțiale prin punctele

$$a = x_0 < x_1 < x_2 < \ldots < x_i < x_{i+1} < \ldots < x_n = b.$$

Să construim pe fiecare interval $[x_i, x_{i+1}]$ $(i = 0, 1, \ldots, \ldots, n - 1)$ un dreptunghi de bază $x_{i+1} - x_i$ și de înălțime $f(x_i)$. Reunind dreptunghiurile astfel obținute, obținem o porțiune P a planului. Aria acestei porțiuni este dată de

$$\text{aria } P = f(a)(x_1 - a) + f(x_1)(x_2 - x_1) + \ldots + f(x_i)(x_{i+1} - x_i) + \ldots + f(x_{n-1})(b - x_{n-1}).$$

Aria lui P diferă de aria lui D. Într-adevăr, dacă privim figura 4.IV, constatăm că P are în plus, față de D, porțiunile hașurate, iar D are în plus, față de P, porțiunile punctate. Este adevărat că se poate întîmpla ca aria totală a porțiunilor hașurate să fie egală cu aria totală a porțiunilor punctate și, datorită acestei compensări, aria lui D să fie egală cu aria lui P.

Aceasta ar fi însă o situație excepțională; în general, această compensare nu se produce. De exemplu, dacă funcția $y = f(x)$ este crescătoare, nu există porțiuni hașurate,

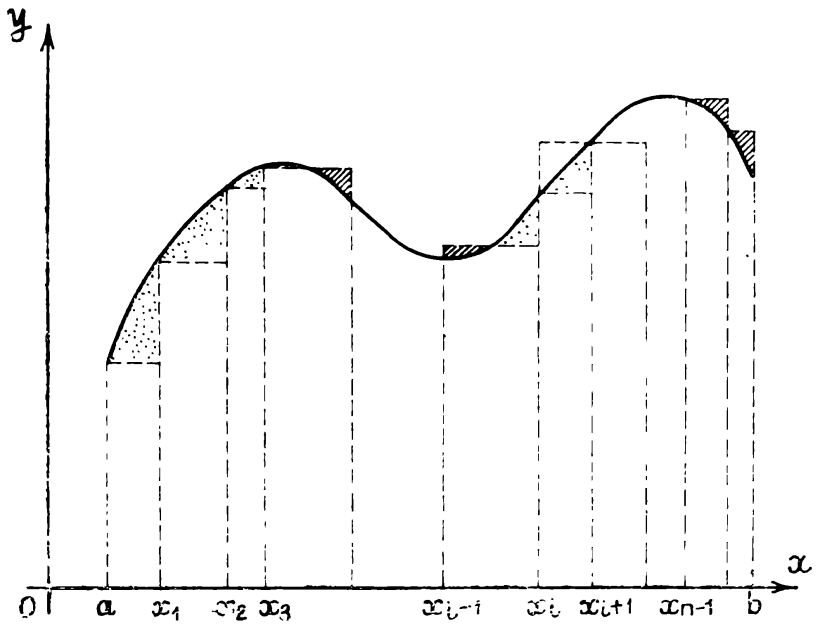


Fig. 4. IV

iar dacă funcția $y = f(x)$ este descrescătoare, lipsesc porțiunile punctate. În primul caz, aria lui P este mai mică decît a lui D; în al doilea caz, aria lui P este mai mare decît aria lui D.

DOUĂ CĂI POSIBILE

Înainte de a merge mai departe, să precizăm o anumită alternativă în legătură cu problema pe care am pus-o. Ne-am propus să determinăm aria porțiunii D. Dacă vrem să facem aceasta în mod exact, nu ne putem opri aici; trebuie să vedem în ce fel se poate elimina eroarea care provine din înlocuirea lui D cu P. Dacă însă ne mulțumim

să determinăm aria lui D nu în mod exact, ci cu o anumită aproximație, de pildă s-o determinăm cu o eroare care să nu depășească o sutime, atunci problema se pune cu totul altfel. Am putea, de exemplu, încerca să alegem punctele $x_1, x_2, \ldots, x_i, \ldots, x_{n-1}$ astfel încît diferența dintre aria totală a porțiunilor haşurate și aria totală a porțiunilor punctate să nu depășească pe 0,01. Cu aceasta însă nu putem spune că am înlocuit problema inițială, a determinării exacte a ariei, cu o problemă mai ușoară. Dificultatea primei probleme provine, de obicei, din structura complicată a curbei C, adică a funcției $y = f(x)$. Se pune deci problema de a înlocui curba C cu alta mai simplă C^*, astfel încît aria porțiunii delimitate de aceasta din urmă să poată fi determinată relativ ușor și, în același timp, să putem determina cît de mare este, în valoare absolută, diferența dintre aria delimitată de C și cea delimitată de C^*. De obicei, în rolul lui C^* se ia o linie poligonală de un anumit tip sau un număr finit de arce de parabolă.

În practică, lucrăm deseori cu valori aproximative. Ne-am putea pune problema dacă nu cumva și pentru determinarea ariei unei porțiuni mai complicate nu ne-am putea mulțumi cu o valoare aproximativă, simplificînd astfel problema. Este clar însă că o evaluare aproximativă a ariei prezintă interes numai dacă sîntem în stare să apreciem de ce ordin de mărime este eroarea comisă. În inginerie se acceptă, în general, o eroare mai mică de o sutime. Tocmai pentru această măsură a erorii este necesar, de multe ori, să cunoaștem și teoria determinării exacte a ariei. După ce cunoști teoria determinării exacte a ariei, te întorci la problema determinării ei aproximative cu forțe noi, cu posibilități noi de investigație. Nu vom insista aici asupra acestei chestiuni, dar importanța ei nu poate fi subapreciată. Ori de cîte ori nu vom fi în stare să determinăm valoarea exactă a ariei — și acest lucru se întîmplă des —, și de multe ori, chiar atunci cînd această determinare este teoretic posibilă, dar este legată de dificultăți de ordin practic, determinarea aproximativă a ariei se impune.

Mai este însă și un alt motiv, mai important decît cel arătat mai sus, pentru care nu ne putem mulțumi numai cu aproximații. Teoria generală luminează calea mai departe, ne duce dincolo de problema concretă de la care am plecat și ne dă posibilitatea să pătrundem mai adînc în cunoașterea relațiilor din natură.

CE SE ÎNTÎMPLĂ CÎND BAZELE DREPTUNGHIURILOR SE MICȘOREAZĂ?

Să închidem paranteza pe care am deschis-o la începutul paragrafului precedent și să ne întoarcem la prima problemă: acea a determinării exacte a ariei. O observație simplă ne va indica drumul pe care trebuie să mergem. Dacă privim figura 4. IV, constatăm că există o legătură foarte strînsă între mărimile porțiunilor hașurate și punctate, pe de o parte, și lungimile bazelor dreptunghiurilor, pe de altă parte. Cu cît aceste baze sînt mai mici, cu atît sînt mai mici și porțiunile prin care diferă D de P. Mai mult decît atît, dacă bazele dreptunghiurilor se micșorează indefinit, atunci atît porțiunile hașurate, cît și cele punctate se micșorează indefinit. Pentru a elimina deci eroarea, trebuie să studiem comportarea expresiei ariei lui P atunci cînd diferențele $x_1 - a$, $x_2 - x_1, \ldots,$ $x_{i+1} - x_i, \ldots, b - x_{n-1}$ se apropie din ce în ce mai mult de zero. Aria porțiunii D va fi tocmai numărul de care se apropie, către care tinde aria porțiunii P atunci cînd bazele dreptunghiurilor se micșorează indefinit. Rămîne numai să vedem care este operația matematică prin care obținem acest număr. Înainte însă de a da răspunsul la această problemă, să mai studiem un exemplu.

CE TREBUIE SĂ ÎNȚELEGEM PRIN LUCRUL MECANIC AL UNEI FORȚE VARIABILE?

Să abordăm acum o problemă de mecanică. Fie o forță f care-și deplasează punctul de aplicație în linie dreaptă, forța fiind îndreptată tot timpul pe direcția de deplasare.

Dacă forța este mereu aceeași, atunci lucrul mecanic pe care-l execută ea este egal cu produsul dintre mărimea forței și mărimea deplasării. Ce se întîmplă însă dacă forța f variază de la punct la punct? Este bine știut că în general tocmai acesta este cazul. De exemplu, mărimea forței de gravitație variază de la punct la punct, deoarece și masa variază de la punct la punct; mărimea unei forțe elastice depinde de distanța față de un anumit punct fix al punctului în care este aplicată forța. În astfel de cazuri, ce trebuie să înțelegem prin lucru mecanic? Este clar că orice încercare de a răspunde la această întrebare trebuie să pornească de la acel caz particular în care răspunsul este cunoscut: cazul în care forța este aceeași în fiecare punct.

Fie $[a, b]$ segmentul de dreaptă de-a lungul căruia forța f își deplasează punctul de aplicație. Fie $f(x)$ mărimea forței f într-un punct oarecare x cuprins între a și b. Așa cum se întîmplă în general în natură, vom admite că forța f variază în mod continuu. În acest fel, este clar că cît diferența $b—a$ este mai mică, cu atît este mai mică și eroarea pe care o comitem presupunînd că forța f este invariabilă. Însă, noi nu dispunem de mărimea $b—a$. Această mărime este dată; modificînd-o, modificăm însăși problema pe care trebuie să o rezolvăm. Vom proceda atunci altfel. Vom împărți intervalul $[a, b]$ în intervalele parțiale prin punctele

$$a = x_o < x_1 < x_2 < \ldots < x_i < x_{i+1} < \ldots < x_n = b.$$

Variația forței $f(x)$ între x_i și x_{i+1} devine oricît de mică dorim, cu condiția ca $x_{i+1} — x_i$ să fie destul de mic. Aceasta înseamnă că pentru $x_i \leqslant x \leqslant x_{i+1}$, presupunînd că forța aplicată în x este nu $f(x)$, ci $f(x_i)$, comitem o eroare egală în valoare absolută cu $| f(x) — f(x_i) |$. Această eroare poate fi făcută oricît de mică, dacă avem grijă să luăm pe x_{i+1} destul de aproape de x_i. Dar o astfel de ipoteză ne reduce problema tocmai la cazul în care forța este constantă. Lucrul mecanic de la x_i la x_{i+1} va fi dat de produsul $f(x_i)(x_{i+1} — x_i)$. Lucrul mecanic total, de la a la b, va fi dat de suma

$$f(a)(x_1—a) + f(x_1)(x_2—x_1) + f(x_2)(x_3—x_2) + \ldots + f(x_{n-1})(b — x_{n-1}).$$

Pentru a elimina eroarea care provine din presupunerea că forța este constantă pe porțiuni, va trebui să studiem comportarea expresiei de mai sus atunci cînd numerele $x_{i+1} - x_i$ se micșorează indefinit. Valoarea exactă a lucrului mecanic al forței $f(x)$, de la a la b, va fi dată de numărul de care se apropie suma de mai sus, atunci cînd diferențele $x_1 - a$, $x_2 - x_1, \ldots, x_{i+1} - x_i, \ldots, b - x_{n-1}$ se apropie de zero. Rămîne numai să vedem, ca și în exemplul precedent, cum exprimăm aceasta într-un mod precis, matematic.

ASEMĂNĂRI ȘI DEOSEBIRI ÎNTRE CELE DOUĂ EXEMPLE DE MAI SUS. NOȚIUNEA DE INTEGRALĂ

Să încercăm să comparăm cele două exemple de mai sus Din punctul de vedere al conținutului proceselor înfățișate, aceste exemple sînt de natură foarte diferită. Primul este de natură geometrică, în timp ce al doilea este luat din fizică. În primul exemplu, $f(x)$ reprezintă înălțimea unui dreptunghi cu vîrful de jos, din stînga, situat în punctul de abscisă x, în al doilea exemplu, $f(x)$ reprezintă forța aplicată în x. În sfîrșit, în primul exemplu suma considerată reprezintă o arie, pe cînd în exemplul al doilea aceeași sumă reprezintă un lucru mecanic.

Cele două procese prezintă totuși o analogie izbitoare. Am putea spune că din punct de vedere formal ele sînt identice. Dacă le golim de conținutul lor specific, adică facem abstracție de acest conținut, constatăm că aceste procese conduc la aceeași schemă, sînt de aceeași formă. Și într-un caz și în celălalt, s-a pornit de la o funcție f definită pe un interval $[a, b]$. S-a împărțit acest interval în intervale parțiale, efectuîndu-se o diviziune — să-i spunem D — a lui $[a, b]$:

$$D = (a = x_0 < x_1 < x_2 < \ldots < x_i < x_{i+1} < \ldots < x_n = b).$$

Acestei diviziuni D i s-a asociat expresia

$$S(D) = f(a)(x_1 - a) + f(x_1)(x_2 - x_1) + \ldots + f(x_{n-1})(b - x_{n-1}) \qquad (1)$$

S-a constatat apoi că în ambele cazuri este necesar să studiem comportarea sumei $S(D)$ date de (1) atunci cînd numerele $x_{i+1} - x_i$ se apropie de zero. Acum a sosit momentul să precizăm această afirmație, să-i dăm un sens matematic.

Să considerăm din nou toate diferențele $x_1 - \ldots$ $\ldots, x_{i+1} - x_i, \ldots b - x_{n-1}$. Să ne îndreptăm atenția asupra celei mai mari dintre aceste diferențe și s-o notăm pe aceasta cu $|D|$ (alte notații: $||D||$ sau $\nu(D)$. Numărul $|D|$ se numește *norma diviziunii* D. Nu este greu de înțeles importanța normei; în loc să spunem de pildă, că fiecare dintre diferențele $x_{i+1} - x_i$ este inferioară numărului 1, va fi suficient să spunem că $|D| < 1$). Din faptul că prin definiție fiecare diferență $x_{i+1} - x_i$ ($i = 0, \ldots, n-1$) este mai mică sau egală cu $|D|$, va rezulta atunci și prima afirmație, anume că toate diferențele $x_{i+1} - x_i$ sînt inferioare lui 1.

Fie acum un șir de diviziuni $D_1, D_2, \ldots, D_p, \ldots$ ale lui $[a, b]$. Să presupunem că șirul $|D_1|, |D_2|, \ldots, |D_p|, \ldots$ tinde la zero. Exemplele considerate anterior ne sugerează că trebuie să studiem comportarea sumei $S(D_p)$ atunci cînd n crește, cu alte cuvinte să ne îndreptăm atenția asupra limitei șirului $S(D_1), S(D_2), \ldots, S(D_p), \ldots$, adică asupra numărului $I = \lim S(D_p)$. Suma $S(D_p)$ se apropie oricît de mult de I, cu condiția ca n să fie suficient de mare. Mai precis, fiind dat un număr $p' > 0$, se poate determina un număr N, astfel încît pentru $p > N$ să avem $|I - S(D_p)| < p'$. Deci, apropierea de I a lui $S(D)$ este condiționată de mărimea (mai bine zis de micimea) normei diviziunii D.

Procesul înfățișat mai sus este un *proces de integrare*, iar numărul I, rezultatul acestui proces, se numește *integrala funcției* f de la a la b și se notează

$$I = \int_a^b f(x)\,dx. \tag{2}$$

De multe ori însă, se înțelege prin *integrală* întregul proces de integrare, împreună cu numărul la care acest proces conduce. Poate, totuși, că n-ar fi rău să diferențiem terminologia referitoare la această chestiune. Este

foarte important să accentuăm caracterul de *proces* al celor înfățișate mai sus.

Semnul $\int$ este o deformare a literei S, inițiala cuvîntului „sumă". El ne amintește că prima etapă în procesul de integrare îl constituie considerarea unor sume. Era o vreme cînd anumite tipuri de integrale se notau chiar cu litera S. Astfel, de pildă, în unele lucrări mai vechi ale matematicianului român Dimitrie Pompeiu poate fi des întîlnită această notație. În prezent, această notație a fost părăsită aproape cu totul. Acest fapt este îndreptățit, deoarece integrala nu se obține prin simpla considerare a unor sume, ci prin supunerea acestor sume unei operații specifice analizei matematice, operația de trecere la limită. Tocmai pentru aceasta, notația integralei trebuia să se deosebească întrucîtva de litera S. Notațiile trebuie să fie cît mai simple, dar și cît mai adecvate obiectului, pentru ca prin intermediul lor să gîndim mai ușor asupra faptelor matematice în cauză. Notația dx amintește de diferențele $x_{i+1} - x_i$, iar $f(x)dx$ de $f(x_i)(x_{i+1} - x_i)$. Desigur că simbolul (2) trebuie luat ca un tot. O componentă oarecare a acestui simbol, ca de exemplu $\int$ sau dx, poate să fie lipsită de sens sau să aibă diverse sensuri. Alăturarea $f(x)dx$ nu trebuie înțeleasă, în general, ca o operație de înmulțire după cum dx nu trebuie privit, în general, ca o diferențială. Nu este însă mai puțin adevărat că în anumite împrejurări, ca de exemplu în problema schimbării de variabilă într-o integrală, dx se comportă exact ca o diferențială, iar $f(x)dx$ are semnificația de produs dintre $f(x)$ și dx. Tocmai aceste împrejurări explică notația întrebuințată.

Semnificația numărului I, adică a integralei, depinde de conținutul specific al procesului de integrare pe care-l efectuăm, adică al procesului prin care, pornind de la $f(x)$, ajungem la I. Astfel, în raport cu primul exemplu, I reprezintă aria porțiunii din plan determinate de graficul lui $f(x)$, axa absciselor și paralele prin a și b la Oy. În raport cu al doilea exemplu, I reprezintă lucrul mecanic pe care-l execută forța $f(x)$ cînd își deplasează punctul de aplicație de la a la b.

Am considerat mai sus doar două exemple, dar ne putem convinge ușor că procedeul prin care se determină spațiul parcurs de un mobil, atunci cînd se cunoaște viteza mobilului în fiecare moment, sau procedeul prin care se determină volumul unui corp de rotație, atunci cînd se cunoaște aria secțiunii perpendiculare pe axa de rotație, ca și multe alte procese din natură conduc, toate, la acel șir de operații pe care le-am descris mai sus și care asociază unei funcții definite în fiecare punct al unui interval un număr real, numit integrala funcției pe acel interval.

Sîntem astfel în posesia unui formalism matematic care sintetizează o varietate extraordinară de procese din natură, reținîndu-le substratul comun, partea lor pur cantitativă, formală. Trebuie, totuși, să observăm că în procesul de integrare a unei funcții există o etapă care, deși clară din punct de vedere teoretic, prezintă dificultăți din punctul de vedere al realizării ei efective. Ne gîndim la acel moment în care trecem de la $S(D)$ la numărul I. Numărul I arată tendința sumelor $S(D)$ atunci cînd norma lui D descrește, I este valoarea de care se apropie $S(D)$ pe măsură ce norma $|D|$ se apropie de zero. Dar numărul I nu ni se dă, urmînd doar ca noi să verificăm că el satisface condiția cerută; *noi* sîntem acei care trebuie să determinăm acest număr. Este adevărat că însăși definiția integralei arată prin ce operație obținem valoarea ei. Trebuie să determinăm limita șirului $S(D_1)$, $S(D_2)$,... $S(D_n)$,... Numai că această operație de determinare a limitei unui astfel de șir este, de cele mai multe ori, chiar în cazul unor funcții simple, legată de dificultăți practice deosebit de mari. Încercați să determinați această limită în cazul unei funcții foarte familiare, cum este $y = \sin x$ de exemplu, considerate pe intervalul [0, 1], și veți vedea că chiar dacă lucrați cu diviziuni formate cu intervale de aceeași mărime, nu veți reuși s-o scoateți la capăt dacă nu stăpîniți bine formulele de trigonometrie.

Or, pentru un astfel de instrument de oglindire matematică a realității, este imperios necesar să găsim mijloace cît mai practice, mai operative de utilizare.

Descoperirea mijlocului de a transforma integrala într-un instrument nu numai virtual, dar și real eficace în descrierea anumitor procese, reprezintă unul din cele mai importante evenimente din istoria matematicilor, poate din întreaga istorie a științelor. Acest eveniment s-a petrecut în urmă cu aproape trei sute de ani. La sfîrșitul secolului al XVII-lea, Newton și Leibniz, independent unul de altul, au descoperit o legătură profundă între noțiunea de integrală și o altă noțiune, care în aparență nu are nici o legătură cu integrala, anume noțiunea de derivată. Este, într-adevăr, surprinzător acest lucru, dacă ne gîndim la modul atît de diferit în care aceste noțiuni au fost definite. Vom vedea însă că de îndată ce pătrundem în esența acestor noțiuni și le înțelegem bine sensul, legătura dintre ele ne apare foarte naturală. Se întîmplă de multe ori ca adevărul științific să aibă la primul contact o aparență paradoxală, să ne surprindă și uneori să vină chiar în contradicție directă cu intuiția imediată. Aceasta datorită faptului că în afară de legăturile exterioare, de suprafață, există între unele fenomene, între unele procese, o legătură lăuntrică, care privește însăși esența acestor fenomene sau procese. O astfel de legătură nu se lasă observată de la prima vedere; pentru a o sesiza nu este suficientă contemplarea pasivă a fenomenelor sau proceselor respective. Însă studiul unui mare număr de procese și separarea, cu grijă, la fiecare dintre ele a formei de conținut, o astfel de operație de abstractizare și sinteză este în măsură să ne conducă la adevăruri de mare valoare științifică.

Pentru a dezvălui această legătură, care conduce imediat la celebra formulă a lui Leibniz și Newton, am putea să reluăm unul dintre exemplele considerate mai sus, cel privitor la arie sau la lucru mecanic. Preferăm însă să folosim un alt exemplu, să pornim de la o altă problemă, care ni se pare mai semnificativă.

Să considerăm o bară materială unidimensională AB. Fie a abscisa lui A și b abscisa lui B. Dacă bara este omogenă, atunci densitatea mijlocie a barei este aceeași pe oricare porțiune a ei. Cu alte cuvinte, pentru orice

porțiune a barei, raportul dintre masa acestei porțiuni și lungimea ei este același. Prin însuși acest fapt, noțiunea de densitate mijlocie este nesemnificativă, inutilă în cazul unei bare omogene.

Să presupunem acum că bara AB nu este omogenă. Aceasta înseamnă că două porțiuni de aceeași lungime pot să fie de mase diferite. Pentru a căpăta o imagine a repartiției masei pe diferitele porțiuni ale barei, sîntem conduși să considerăm raportul dintre masa unei porțiuni și lungimea ei. Acest raport ne indică densitatea mijlocie a barei pe porțiunea respectivă, adică masa ce se găsește în medie, pe unitatea de lungime, pe porțiunea respectivă. De exemplu, dacă notăm cu $F(x)$ masa ce se găsește între A și punctul X de abscisă x, considerînd două puncte X' și X'' ale barei, de abscise, respectiv, x' și x'', densitatea mijlocie a porțiunii $X'X''$ va fi dată de raportul

$$\frac{F(x'') - F(x')}{x'' - x'}. \tag{3}$$

Acest raport nu ne spune însă nimic despre modul de repartiție a masei pe diferitele porțiuni ale intervalului $[x', x'']$. Sîntem astfel conduși, pentru a căpăta o imagine a acestei repartiții, să cercetăm densitatea mijlocie a barei pe porțiuni din ce în ce mai mici. O dată porniți pe această cale, nu ne mai putem opri la jumătatea drumului; trebuie să cercetăm comportarea raportului (3) atunci cînd diferența $x'' - x'$ se micșorează indefinit. În felul acesta, ajungem să obținem o imagine locală, punctuală, a repartiției masei. Iată ce trebuie să înțelegem prin aceasta.

Fie X_0 un punct cuprins între A și B. Să notăm cu x_0 abscisa lui X_0. Raportul

$$\varphi(h) = \frac{F(x_0 + h) - F(x_0 - h)}{2h}$$

reprezintă densitatea mijlocie a barei pe intervalul $[x_0 - h, x_0 + h]$ cu centrul în x_0. Să presupunem că $\varphi(h)$ admite o limită finită cînd h tinde la zero. Cu alte cuvinte, există un număr α, astfel încît pentru orice șir $h_1, h_2, \ldots, h_n, \ldots$ tinzînd la zero, șirul $\varphi(h_1), \varphi(h_2), \ldots, \varphi(h_n), \ldots$ tinde la α. Valoarea α a limitei astfel obținute este, prin definiție, densitatea barei AB în punctul X_0.

Este evidentă deosebirea dintre densitatea mijlocie și densitatea într-un punct. Prima se referă la un întreg interval, are deci un caracter global, integral, și se obține prin operații de scădere și împărțire; a doua se referă la un punct, are deci un caracter local, punctual, și se obține printr-o operație, un proces de trecere la limită.

Dar, între funcția $F(x)$, care reprezintă masa ce se găsește între A și X, pe de o parte, și densitatea punctuală a barei, pe de altă parte, există o legătură strînsă; derivata funcției $F(x)$ într-un punct x_0 este tocmai densitatea barei în punctul x_0. Într-adevăr, avem

$$\frac{F(x_0+h)-F(x_0-h)}{2h}=\frac{1}{2}\left[\frac{F(x_0+h)-F(x_0)}{h}-\frac{F(x_0)-F(x_0-h)}{h}\right],$$

și deoarece, punînd $-h=k$, obținem

$$\lim_{h\to 0}\frac{F(x_0)-F(x_0-h)}{h}=\lim_{k\to 0}\frac{F(x_0+k)-F(x_0)}{k}=F'(x_0),$$

deducem

$$\lim_{h\to 0}\varphi(h)=\frac{1}{2}\left[\lim_{h\to 0}\frac{F(x_0+h)-F(x_0)}{h}+\lim_{k\to 0}\frac{F(x_0)-F(x_0-h)}{h}\right]=$$
$$=\frac{1}{2}\left[F'(x_0)+F'(x_0)\right]=F'(x_0).$$

Cunoscînd deci masa unei bare, putem afla densitatea punctuală a barei. Trecerea de la una la cealaltă se face printr-un proces de derivare. Masa este o caracteristică globală a barei, iar densitatea este o caracteristică locală, punctuală.

Prin considerații analoge, binecunoscute, cunoscînd spațiul parcurs de un mobil în orice interval de timp (deci caracteristica integrală, în mare, a unei mișcări), putem determina, printr-un proces de derivare, viteza la un moment oarecare a mobilului (deci caracteristica momentană, instantanee, a mișcării). Ne apare astfel semnificația comună tuturor proceselor de derivare: trecerea de la descrierea în mare a unui fenomen, la descrierea lui locală sau momentană.

Să încercăm acum să punem problema invers. Cunoscînd densitatea în fiecare punct a unei bare neomogene, să determinăm masa barei.

Deoarece densitatea variază în mod continuu, presupunerea că bara ar fi omogenă este cu atît mai apropiată de realitate, comportă o eroare cu atît mai mică, cu cît bara este mai mică. Însă, neputînd modifica lungimea barei, vom presupune că bara este omogenă pe porțiuni, cu alte cuvinte există o diviziune D a intervalului $[a, b]$ în intervale parțiale $a = x_0 < x_1 < \ldots < x_i < x_{i+1} < \ldots < x_n = b$, astfel încît porțiunea $[x_i, x_{i+1}]$ este omogenă, oricare ar fi $i = 0, 1, \ldots, n - 1$. Aceasta înseamnă că în fiecare punct al intervalului $[x_i, x_{i+1}]$ densitatea este aceeași, de pildă egală cu valoarea ei în punctul x_i. În acest caz, masa porțiunii $[x_i, x_{i+1}]$ este egală cu produsul dintre lungimea acestei porțiuni și densitatea ei, adică cu $f(x_i)(x_{i+1} - x_i)$ (prin $f(x)$ notăm densitatea barei în punctul x). Masa (aproximativă) a barei este

$$S(D) = f(a)(x_1 - a) + f(x_1)(x_2 - x_1) + \ldots + f(x_{n-1})(b - x_{n-1})$$

și această masă aproximativă este cu atît mai apropiată de masa exactă a barei, cu cît lungimile intervalelor diviziunii D sînt mai mici. Sîntem deci conduși, pentru a elimina eroarea provenită din aproximarea pe care am făcut-o, să considerăm un șir de diviziuni D_n, pentru care $\lim_{n\to\infty} |D_n| = 0$. Masa exactă a barei va fi dată de limita șirului

$$S(D_1),\ S(D_2),\ \ldots,\ S(D_n),\ldots .$$

După cum vedem, operațiile prin care, pornind de la densitatea punctuală a barei, determinăm masa ei, sînt, din punctul de vedere al formei lor, exact aceleași pe care le-am folosit pentru a determina aria porțiunii din plan delimitată de o curbă, cînd se cunoaște ordonata fiecărui punct al curbei, sau pentru a determina lucrul mecanic al unei forțe atunci cînd se cunoaște mărimea forței în fiecare punct. Deci, cunoscînd densitatea, masa se obține printr-un proces de integrare. Dacă $f(x)$ este densitatea barei în punctul x, atunci masa barei este

$$\int_a^b f(x)\mathrm{d}x.$$

O trăsătură comună tuturor exemplelor considerate este trecerea de la descrierea locală punctuală sau instantanee a unui fenomen la descrierea lui în mare, adică tocmai procesul invers derivării, care, după cum am văzut mai sus, ne dă posibilitatea să trecem de la descrierea globală a unui fenomen la descrierea lui locală.

Această trăsătură a procesului de derivare apare nu numai cînd se definește densitatea; ea aparține oricărui proces de derivare, și semnificația derivării stă tocmai în această trăsătură. Deci, integrarea și derivarea sînt două procese opuse, inverse unul altuia, așa cum sînt opuse operațiile de adunare și scădere sau cele de înmulțire și împărțire. Iată un tablou care redă acest aspect al derivării și integrării:

Caracteristica locală	integrare ⇄ derivare	Caracteristica globală
Ordonata punctului de abscisă x de pe graficul funcției $f(x) > 0$ definită pe $[a, b]$.	integrare ⇄ derivare	Aria porțiunii din plan delimitată de dreptele $x = a$, $x = x$ și curba $y = f(x)$.
Viteza unui mobil la fiecare moment t $(a \leqslant t \leqslant b)$.	integrare ⇄ derivare	Spațiul parcurs de mobil în intervalul de timp de la momentul $t=a$ la momentul $t=t$ $(a \leqslant t \leqslant b)$.
Densitatea punctuală a unei bare de abscise extreme a și b.	integrare ⇄ derivare	Masa totală a porțiunii $[a, x]$ din bară.
Aria secțiunii perpendiculare pe axa de rotație într-un corp descris prin rotația de 360° a curbei $y = f(x)$ $(a \leqslant x \leqslant b, f(x) > 0)$ în jurul axei Ox.	integrare ⇄ derivare	Volumul corpului de rotație corespunzător segmentului $[a, x]$, $(a \leqslant x \leqslant b)$.
Densitatea punctuală a unei sarcini electrice liniare distribuite pe $[a, b]$.	integrare ⇄ derivare	Sarcina electrică totală pe porțiunea $[a, x]$, unde $a \leqslant x \leqslant b$.
Mărimea forței $f(x)$ pentru fiecare x $(a \leqslant x \leqslant b)$.	integrare ⇄ derivare	Lucrul mecanic total al forței f de la a la x $(a \leqslant x \leqslant b)$.

Sub acest raport, al trecerii de la aspectul local la aspectul global al unui fenomen, noțiunea de integrală constituie punctul de plecare al uneia dintre cele mai importante ramuri ale analizei matematice: teoria ecuațiilor diferențiale. A integra o funcție $f(x)$ înseamnă a găsi o altă funcție $F(x)$ a cărei derivată este $f(x)$, adică $F'(x) = f(x)$. O astfel de ecuație, în care necunoscuta este o funcție, anume $F(x)$, și această funcție intervine prin derivata ei, se numește o ecuație diferențială. Dar legătura între aspectul local și cel global al unui fenomen nu este dată totdeauna sub această formă simplă. Nenumărate probleme de geometrie, de fizică, de chimie etc. conduc la legături mai complicate între aspectele locale și cele globale, adică la ecuații diferențiale mai complicate decît cea de mai sus. Studiul acestor ecuații este deosebit de important pentru descrierea matematică a fenomenelor naturii, și teoria lor este una dintre cele mai delicate ale matematicii moderne.

În cadrul școlii matematice românești există bogate tradiții privind studiul ecuațiilor diferențiale. Amintim aici teza de doctorat a lui Simion Stoilow, contribuțiile lui Nicolae Ciorănescu, ale academicienilor Miron Nicolescu, Grigore Moisil și Nicolae Teodorescu și ale profesorului Ciprian Foiaș (în domeniul ecuațiilor diferențiale cu derivate parțiale), precum și teza de doctorat a lui Anton Davidoglu și contribuțiile lui V. M. Popov, membru corespondent al Academiei, și ale profesorilor I. Barbălat, C. Corduneanu și A. Halanay (în domeniul ecuațiilor diferențiale ordinare).

MOMENTUL CAUCHY

În prezentarea de mai sus, am lăsat deoparte unele aspecte existențiale, calitative sau care țin aproape exclusiv de dezvoltarea internă a teoriei integralei. Aceste aspecte au fost studiate abia atunci cînd analiza matematică a atins un grad mai înalt de dezvoltare, anume începînd cu a doua jumătate a secolului trecut. Am urmărit doar să explicăm geneza uneia dintre cele mai importante

noțiuni ale matematicii, rădăcinile ei adînci în realitate. În cele ce urmează vom urmări mai departe evoluția ideii de integrală, arătînd ce a determinat, de fiecare dată, lărgirea vechiului cadru, punînd accentul de fiecare dată pe aspectele intuitive privind legătura teoriei integralei cu ceea ce este în afara ei.

La începutul secolului trecut, Cauchy era preocupat de integrarea funcțiilor continue. În acest scop, el a folosit sume de tipul

$$\sigma_\Delta(f) = \sum_{i=0}^{n-1} f(x_i)(x_{i+1} - x_i) \qquad (4)$$

(unde Δ este diviziunea $a = x_0 < x_1 < \ldots < x_i < x_{i+1} < \ldots < x_n = b$, iar f este o funcție continuă pe $[a, b]$) și a demonstrat că există un număr natural I cu proprietatea următoare: pentru orice $\varepsilon > 0$ există $\eta > 0$, astfel încît oricare ar fi diviziunea Δ de normă inferioară lui η, avem $|\sigma_\Delta(f) - I| < \varepsilon$. Sumele de tipul (4) se numesc sume Cauchy, integrabilitatea în sensul definit mai sus se numește integrabilitate Cauchy, iar numărul I se numește integrala Cauchy a funcției f pe $[a, b]$. Rezultă că orice funcție continuă este integrabilă Cauchy. Pe de altă parte, se știe că o funcție continuă pe un interval admite primitivă pe acest interval, adică există o funcție derivabilă. G, astfel încît $G' = g$ pe I. Fie F o primitivă a funcției f. O teoremă celebră a lui Leibniz și Newton — teoremă care, într-un anumit caz particular, a fost prezentată în paragraful precedent — ne asigură că valoarea I a integralei lui f pe $[a, b]$ este tocmai diferența $F(b) - F(a)$. Cu alte cuvinte, atîta timp cît ne restrîngem la clasa funcțiilor continue, procesul de derivare și cel de integrare (așa cum a fost conceput de Cauchy) apar efectiv ca două procese duale, în concordanță cu schema generală a proceselor de derivare și integrare, schemă descrisă în paragraful precedent.

MOMENTUL RIEMANN

Dar dezvoltarea analizei matematice a condus, încă din prima jumătate a secolului trecut — în special prin teoria seriilor trigonometrice — la considerarea funcțiilor

discontinue. Serii trigonometrice dintre cele mai simple au ca sumă funcții discontinue. Este suficient să ne amintim că însăși funcția lui Dirichlet, egală cu 1 pentru x rațional și cu 0 pentru x irațional, este suma unei serii trigonometrice și tocmai acest fapt l-a determinat pe Dirichlet s-o considere.

Dar, așa cum în studiul seriilor de puteri un rol fundamental îl joacă derivata, în studiul seriilor trigonometrice nu ne putem dispensa de noțiunea de integrală. Coeficienții unei serii de puteri se exprimă cu ajutorul sumei seriei, folosind operația de derivare. Coeficienții unei serii trigonometrice convergente pe un interval J se exprimă uneori cu ajutorul sumei acestei serii, folosind operația de integrare pe J. Aceasta presupune deci că funcția-sumă se poate integra. Trebuie deci să imaginăm un procedeu de integrare care să se aplice și funcțiilor discontinue, deoarece suma unei serii trigonometrice nu este totdeauna o funcție continuă.

Ideea lui Riemann a fost aceea de a considera procedeul utilizat de Cauchy în integrarea funcțiilor continue și de a stabili limitele de aplicabilitate a acestui procedeu. Deci, în timp ce Cauchy își dădea o funcție continuă arbitrară și-și propunea s-o integreze printr-un procedeu fixat, Riemann își dă o funcție despre care nu mai presupune că este continuă, ci doar că i se poate aplica procedeul de integrare propus de Cauchy. O astfel de funcție primește numele de funcție integrabilă Riemann, iar integrala ei este integrala Riemann (a se vedea și sfîrșitul capitolului I') Deci în timp ce Cauchy este preocupat de calculul integralei unei funcții continue, Riemann își îndreaptă atenția principală asupra ideii de integrabilitate. În fapt, Riemann nu consideră chiar sume de tipul (4), ci niște sume de o formă mai generală și anume:

$$\sigma_{\Delta}(f,\ \xi_i) = \sum_{i=0}^{n-1} f(\xi_i)(x_{i+1} - x_i), \tag{5}$$

unde $x_i \leqslant \xi_i \leqslant x_{i+1}$, pentru $0 \leqslant i \leqslant n-1$. O funcție f este deci integrabilă Riemann pe $[a, b]$, dacă există un număr I cu următoarea proprietate: oricare ar fi $\varepsilon > 0$, există $\eta > 0$, astfel încît dacă norma diviziunii Δ este mai

mică decît η, atunci $|I - \sigma_\Delta(f, \xi_i)| < \varepsilon$, oricare ar fi valorile ξ_i pentru care $x_i \leqslant \xi_i \leqslant x_{i+1}$ $(0 \leqslant i \leqslant n-1)$. Numărul I se numește integrala Riemann a funcției f pe $[a, b]$.

După Riemann, efortul matematicienilor s-a îndreptat spre caracterizarea structurală a funcțiilor integrabile Riemann, pentru a se stabili, în felul acesta, cît de mult depășește clasa funcțiilor integrabile Riemann pe aceea a funcțiilor continue. Punctul culminant al acestui efort este marcat de un rezultat — astăzi clasic — datorit lui Lebesgue și care constituie una dintre cele mai frumoase teoreme ale analizei matematice. Lebesgue a demonstrat că o funcție mărginită este integrabilă Riemann pe $[a, b]$, dacă și numai dacă ea este continuă aproape peste tot pe $[a, b]$. Amintim că o proprietate punctuală are loc aproape peste tot, dacă punctele în care ea nu se verifică alcătuiesc o mulțime de măsură nulă, adică (a se vedea și capitolul I) o mulțime A, cu proprietatea că pentru orice număr $\varepsilon > 0$ se poate găsi o familie cel mult numărabilă de intervale de lungime totală inferioară lui ε, care acoperă mulțimea A. Este destul să observăm că există mulțimi infinite (chiar nenumărabile), care sînt de măsură nulă, pentru ca să ne dăm seama că clasa funcțiilor integrabile Riemann este mult mai largă decît clasa funcțiilor continue. În particular, orice funcție cu variație mărginită este integrabilă Riemann.

Oricărei funcții f integrabile Riemann pe $[-\pi, \pi]$ i se asociază o serie Fourier-Riemann, adică o serie trigonometrică

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx)$$

ai cărei coeficienți a_n și b_n sînt dați de expresiile

$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos nx \, dx, \quad b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin nx \, dx,$$

unde integrala este considerată în sensul lui Riemann. În particular, oricărei funcții cu variație mărginită pe $[-\pi, \pi]$ i se asociază o serie Fourier-Riemann pe $[-\pi, \pi]$. O teoremă clasică a lui Jordan afirmă că această serie

este convergentă în fiecare punct și admite ca sumă pe $f(x)$ în orice punct x în care f este continuă; în plus, convergența seriei este uniformă pe orice subinterval compact în care funcția este continuă.

„DIVORȚUL" DINTRE PRIMITIVĂ ȘI INTEGRALĂ

Am văzut că pentru o funcție f continuă pe $[a, b]$, integrala de la a la b se obține făcînd diferența valorilor primitivei lui f în punctele b și a. Tocmai în această proprietate stă întreaga fecunditate a calculului diferențial și integral clasic. Extensiunea adusă de Riemann ideii de integrabilitate, lărgind considerabil clasa funcțiilor care se pot integra, pune în același timp o problemă gravă: o funcție integrabilă Riemann admite totdeauna primitivă? În cazul afirmativ, mai este în vigoare aici o formulă de tipul Leibniz-Newton? Trebuie să spunem că răspunsul la prima întrebare este negativ, deci problema a doua nu se mai pune. Este suficient să ne gîndim la funcția lui Riemann, definită pe $[a, b]$ $(a > 0)$ în felul următor:

$$f(x) = \begin{cases} 0, \text{ dacă } x \text{ este irațional} \\ \frac{1}{q}, \text{ dacă } x = \frac{p}{q}, \text{ unde } p \text{ și } q \text{ sînt} \\ \text{numere întregi, } q > 0, \\ \text{iar } p \text{ și } q \text{ sînt prime între ele.} \end{cases}$$

Această funcție este evident mărginită, iar mulțimea punctelor ei de discontinuitate coincide cu mulțimea punctelor raționale din $[a, b]$, deci este o mulțime numărabilă. Dar o mulțime numărabilă este de măsură nulă deci, în baza teoremei fundamentale a lui Lebesgue, semnalată în paragraful precedent, funcția considerată este integrabilă Riemann pe $[a, b]$. Pe de altă parte, această funcție admite în fiecare punct rational o discontinuitate de prima specie (avem $f(x - 0) = f(x + 0) = 0$ pentru orice x din $[a, b]$), deci f nu admite primitivă (conform unei teoreme care afirmă că discontinuitățile unei funcții derivate nu pot fi de prima specie).

Acest exemplu arată cît de mult se poate deosebi o funcție integrabilă Riemann de o funcție care admite primitivă. Există însă o situație de compromis: Dacă o funcție f, integrabilă Riemann pe $[a, b]$, admite primitivă — fie ea F — pe $[a, b]$, atunci

$$\int_a^b f(x)\mathrm{d}x = F(b) - F(a).$$

Pentru a ne asigura că această teoremă constituie un progres real în raport cu formula lui Leibniz-Newton pentru funcții continue, este necesar să dăm un exemplu de funcție integrabilă Riemann pe $[a, b]$ care nu este continuă pe $[a, b]$, dar admite primitivă pe $[a, b]$. Un astfel de exemplu este funcția

$$\varphi(x) = \begin{cases} 2x\sin\frac{1}{x} - \cos\frac{1}{x}, & \text{dacă } x \neq 0 \\ 0, & \text{dacă } x = 0 \end{cases}$$

Această funcție este mărginită și admite un singur punct de discontinuitate (în origine), deci este integrabilă Riemann pe orice interval compact. În același timp φ admite primitivă pe orice interval; funcția F dată de

$$F(x) = \begin{cases} x^2\sin\frac{1}{x}, & \text{dacă } x \neq 0 \\ 0, & \text{dacă } x = 0 \end{cases}$$

este o primitivă a funcției φ.

Ținînd seamă că există derivate discontinue (funcția φ de mai sus constituie un astfel de exemplu), este natural să ne întrebăm dacă există derivate neintegrabile Riemann. Este de la sine înțeles că această problemă devine interesantă numai dacă ne restrîngem la derivate mărginite, deoarece o derivată nemărginită pe un interval compact nu va avea nici o șansă să fie integrabilă Riemann pe acest interval. Primul exemplu de derivată mărginită, neintegrabilă Riemann, a fost dat la sfîrșitul secolului trecut de către matematicianul italian Vito Volterra. Nu vom reproduce aici acest exemplu, deoarece un exemplu

mai elementar și mai puternic este furnizat de o construcție imaginată de către matematicianul român Dimitrie Pompeiu. Fie

$$f(x) = \sum_{n=1}^{\infty} \frac{1}{2^n} (x - r_n)^{\frac{1}{3}}, \qquad (6)$$

unde $\{r_n\}$ este un șir format cu numerele raționale din intervalul compact $[a, b]$. Se arată relativ ușor că seria din membrul al doilea din (6) este convergentă pe $[a, b]$, iar funcția f este continuă și strict crescătoare pe $[a, b]$. Rezultă că f este inversabilă pe $[a, b]$, iar inversa ei, fie ea g, este continuă și strict crescătoare pe $[f(a), f(b)]$. Mai mult decît atît, se arată că g este derivabilă pe $[f(a), f(b)]$, iar derivata g', deși mărginită, nu este integrabilă Riemann pe nici un subinterval compact al lui $[f(a), f(b)]$. (Exemplul lui Volterra este acela al unei derivate mărginite h, neintegrabile Riemann pe $[a, b]$, dar orice interval conținut în $[a, b]$ conține un subinterval compact, pe care h este integrabilă Riemann.) Recent, un matematician american, Andrew Bruckner, reluînd și continuînd cercetările lui Pompeiu, a reușit să dea un exemplu de derivată mărginită, discontinuă aproape peste tot. Acest exemplu arată cît de mult se poate deosebi o derivată mărginită de o derivată integrabilă Riemann (aceasta din urmă este obligatoriu continuă aproape peste tot).

DOUĂ IDEI ALE LUI LEBESGUE

Am adunat destule probe pentru ca să putem afirma că analiza matematică are nevoie de o integrală mai generală decît aceea în sensul lui Riemann. Să mai semnalăm faptul că există serii trigonometrice convergente, a căror sumă nu este o funcție integrabilă Riemann pe intervalul de convergență. În sfîrșit, să mai amintim faptul că și problema calculului lungimii unei curbe rectificabile a condus la necesitatea de a se considera o integrală mai generală decît integrala Riemann (vezi capitolul III). O astfel de integrală mai generală, care avea să dea răspunsuri cel puțin parțiale la toate problemele în fața cărora inte-

grala Riemann se dovedea neputincioasă, avea să fie introdusă la începutul secolului nostru de către matematicianul francez Henri Lebesgue. Integrala introdusă de către Lebesgue constituie cel mai de seamă eveniment din întreaga dezvoltare a teoriei integralei de la Riemann și pînă în zilele noastre. Însă, înainte de a o prezenta, să încercăm să reconstituim, măcar cu aproximație, mersul gîndirii lui Lebesgue.

După cum se poate constata din capitolul III, Lebesgue obișnuia să se inspire în invențiile sale matematice din cele mai simple intuiții pe care le oferă realitatea imediată. Observînd cu atenție procedeele utilizate în matematica elementară, spiritul critic al lui Lebesgue îndepărta cu curaj orice element neesențial, chiar dacă acesta era intrat atît de mult în obișnuință, încît nimeni nu se mai întreba asupra rostului său.

Să ne întoarcem, pentru un moment, la procedeul de integrare folosit de Riemann. Fiind dată o funcție reală f definită pe $[a, b]$, Riemann asociază fiecarei diviziuni $\Delta = (a = x_0 < x_1 < \ldots < x_i < x_{i+1} < \ldots < x_n = b)$ expresiile de forma

$$\sigma_\Delta (f, \xi_i) = \sum_{i=0}^{n-1} f(\xi_i)(x_{i+1} - x_i). \qquad (7)$$

Din punct de vedere geometric o sumă Riemann (adică o sumă de forma (7)) reprezintă aria unui anumit domeniu poligonal $D_\Delta(f, \xi_i)$, care se obține printr-o reuniune finită de domenii dreptunghiulare D_i $(i = 0, 1, \ldots, n - 1)$, baza lui D_i fiind de mărime $x_{i+1} - x_i$, iar înălțimea fiind egală cu $f(\xi_i)$. Procesul de integrare riemanniană revine, din punct de vedere geometric, la o aproximare a ariei porțiunii din plan delimitate de axa absciselor, dreptele $x = a$, $x = b$ și graficul funcției considerate prin ariile domeniilor poligonale $D_\Delta(f, \xi_i)$ (vezi fig. 4IV). În fapt, integrabilitatea Riemann a funcției f revine la existența și finitudinea limitei

$$\lim_{\nu(\Delta_p)\to 0} \text{aria } D_{\Delta_p}(f, \xi_i^p)$$

și la independența acestei limite atît de șirul de diviziuni Δ_p de normă tinzînd la zero, cît și de valorile intermediare ξ_i^p $(0 \leqslant i \leqslant n-1)$.

După cum vedem, în procesul de integrare riemanniană un rol central îl au diviziunile Δ ale intervalului de definiție a funcției f. Dar o diviziune Δ revine la o partiție finită a intervalului $[a, b]$ în subintervale, deoarece orice diviziune Δ induce o astfel de partiție și invers. Deci punctele din $[a, b]$ sînt clasificate aici luînd drept criteriu ordinea lor de la stînga la dreapta, pe intervalul $[a, b]$. De exemplu, mulțimea de rang i a acestei partiții este alcătuită din punctele din $[a, b]$ care nu sînt la stînga punctului de abscisă x_i, dar se află la stînga punctului de abscisă x_{i+1}.

Lebesgue a observat că acest criteriu al ordinii de la stînga la dreapta (adică al ordinii induse de mărimea abscisei) este un rezultat al rutinei, el nefiind cerut de esența problemei integrării. În fapt, adoptarea acestui criteriu este chiar nepractică, așa cum nepractic este să numărăm o sumă de bani în ordinea întîmplătoare în care monezile sînt scoase din buzunar. Numărătoarea devine mult mai simplă dacă grupăm în prealabil, monezile punînd laolaltă pe cele de 1 leu, de 3 lei, de 5 lei și așa mai departe. Să observăm însă că în timp ce numărătoarea făcută la întîmplare avea un caracter mai elementar, implicînd doar operații de adunare, numărătoarea după a doua metodă, a grupării monezilor după valoarea lor, este ceva mai puțin elementară, implicînd și operații de înmulțire.

Analogia de mai sus, imaginată chiar de Lebesgue, este de natură să ne călăuzească pașii mai departe. Ordinea întîmplătoare în care monezile sînt scoase din buzunar este ordinea — întîmplătoare, neesențială pentru problema integrării — de la stînga la dreapta, a punctelor din $[a, b]$. Gruparea monezilor după valoarea lor ne sugerează gruparea punctelor din $[a, b]$ după valorile corespunzătoare ale funcției f. Deoarece o funcție ia în general o infinitate de valori, iar noi dorim ca numărul grupelor pe care le formăm să fie finit, vom proceda în modul următor. Presupunînd, pentru simplificare, că funcția f este mărginită, vom

considera un interval $[A,B]$ cu proprietatea $A<m<M<B$, unde m și M sînt marginile funcției f pe $[a, b]$. Vom considera apoi o diviziune $\delta = (m = y_0<y_1< \ldots <y_i<y_{i+1}< < \ldots < y_m = M)$ și vom asocia acestei diviziuni șirul de mulțimi $E_0, E_1, \ldots, E_i, E_{i+1}, \ldots E_{m-1}$, unde mulțimea E_i conține exact acele puncte x din $[a, b]$, pentru care $y_i \leqslant f(x) < y_{i+1}$. Mulțimile E_i $(0 \leqslant i < m)$ sînt disjuncte două cîte două, iar reuniunea lor este tocmai intervalul $[a, b]$, deci ele definesc o partiție a lui $[a, b]$. Dar, așa cum numărătoarea banilor devine mai practică sacrificînd ceva din caracterul elementar al operațiilor, observăm că în procedeul de mai sus natura mulțimilor E_i este, în general, mai complicată decît a mulțimilor corespunzătoare din procedeul lui Riemann; într-adevăr, acestea din urmă sînt totdeauna intervale (anume, intervalele $[x_0, x_1)$, $[x_1, x_2)\ldots$, $[x_i, x_{i+1})$, $\ldots$, $[x_{n-2}, x_{n-1})$ și $[x_{n-1}, x_n]$), în timp ce mulțimile E_i se complică o dată cu funcția f.

Să vedem acum în ce fel se poate valorifica ideea de mai sus în generalizarea și ameliorarea procesului de integrare. Lebesgue și-a dat seama — și aceasta este a doua mare idee pe care acest savant a introdus-o în teoria integralei — că teoria integralei trebuie precedată de o teorie generală a măsurii mulțimilor, o teorie care să ne învețe să dăm un sens conceptului de măsură pentru mulțimi cît mai generale, așa cum matematica elementară ne învață să definim lungimea unui interval, aria unui domeniu poligonal și volumul unui domeniu poliedral.

Bineînțeles că o astfel de teorie generală a măsurii va trebui să satisfacă anumite cerințe logice și intuitive. De pildă: să furnizeze pentru intervale, domenii poligonale sau poliedrale aceeași măsură pe care o furnizează și matematica elementară; fiind dată o descompunere a unei mulțimi măsurabile A într-un număr finit sau numărabil d emulțimi disjuncte A_i, de asemenea măsurabile, măsura lui A să fie egală cu suma măsurilor mulțimilor A_i (proprietatea de aditivitate).

Despre mulțimile cărora li se poate atribui o măsură (vom vedea în paragraful următor cum anume se poate construi o teorie a măsurii) vom spune că sînt măsurabile. Despre o funcție f pentru care mulțimile $\{x; \alpha \leqslant f(x) < \beta\}$ sînt mă-

surabile, oricare ar fi numerele α și β (inclusiv cazurile $\alpha = -\infty$, $\beta = +\infty$), vom spune că este o funcție măsurabilă. Această definiție a măsurabilității unei funcții se inspiră, după cum se poate vedea ușor, din natura mulțimilor E_i de mai sus; într-adevăr, avem $E_i = \{x; y_i \leqslant f(x) < y_{i+1}\}$, deci funcția f este măsurabilă, atunci putem atribui mulțimilor E_i o măsură, fie aceasta $m(E_i)$. Acum este suficient să formăm sumele

$$s_\delta(f) = \sum_{i=0}^{m-1} y_i m(E_i) \tag{8}$$

$$S_\delta(f) = \sum_{i=0}^{m-1} y_{i+1} m(E_i) \tag{9}$$

sume care constituie un analog al sumelor Darboux din teoria integralei Riemann (a se vedea capitolul III). Le-am putea numi sume Darboux-Lebesgue; suma (8) este o sumă Darboux-Lebesgue inferioară, iar suma (9) este o sumă Darboux -Lebesgue superioară. Ca și la sumele Darboux obișnuite, avem și aici inegalitatea

$$s_\delta(f) \leqslant S_\delta(f).$$

Continuînd această analogie cu sumele Darboux obișnuite și observînd că nici o sumă $s_{\delta'}$ nu poate întrece o sumă $S_{\delta''}$ (exercițiu!), vom spune că funcția f este integrabilă Lebesgue pe $[a, b]$ dacă marginea superioară a sumelor Darboux-Lebesgue inferioare este egală cu marginea inferioară a sumelor Darboux-Lebesgue superioare. Putem da de pe acum o teoremă care arată puterea și eleganța noului tip de integrabilitate: orice funcție mărginită și măsurabilă pe $[a, b]$ este integrabilă Lebesgue pe $[a, b]$. Într-adevăr, avem

$$S_\delta(f) - s_\delta(f) = \sum_{i=0}^{m-1} (y_{i+1} - y_i) m(E_i)$$

deci, pentru diviziuni δ cu norma mai mică decît $\varepsilon > 0$, avem

$$S_\delta(f) - s_\delta(f) < \varepsilon \sum_{i=0}^{m-1} m(E_i).$$

Însă, ținînd seama de proprietatea de aditivitate a măsurii, avem

$$\sum_{i=0}^{m-1} m(E_i) = b - a,$$

deci

$$S_\delta(f) - s_\delta(f) < \varepsilon(b - a),$$

oricare ar fi diviziunea δ, de normă mai mică decît ε, a intervalului $[A, B]$. Așadar

$$\sup_\delta s_\delta(f) = \inf_\delta S(f)$$

și funcția f este integrabilă Lebesgue pe $[a, b]$.

Însă succesul procedeului de mai sus depinde de posibilitatea de a defini o noțiune de măsură în raport cu care clasa funcțiilor măsurabile să conțină o parte cît mai mare dintre funcțiile uzuale și, bineînțeles, să conțină toate funcțiile integrabile Riemann. Această problemă a fost rezolvată la începutul secolului nostru de către Lebesgue, și ei îi dedicăm paragraful următor.

MĂSURA LEBESGUE

Pentru comoditate, ne vom restrînge la mulțimi ale dreptei numerice, dar cititorul își va da ușor seama că toate considerațiile se extind la mulțimi din plan sau din spațiu.

Matematica elementară ne învață să atribuim o lungime oricărui interval, o arie oricărui domeniu poligonal, un volum oricărui corp poliedral. Vom presupune aceste lucruri cunoscute, totuși observăm că și aici sînt unele aspecte delicate, peste care un începător în matematică trece fără a manifesta prea multă exigență. Vom ilustra această situație prin două exemple.

Dacă definim aria unui triunghi ca jumătate din produsul unei laturi prin înălțimea corespunzătoare, atunci pentru a legitima definiția pusă, este necesar să arătăm că produsul dintre o latură a unui triunghi și înălțimea corespunzătoare nu depinde de latura considerată. Mai departe,

definind aria unui poligon ca sumă a ariilor triunghiurilor în care se descompune poligonul considerat, este necesar să arătăm că valoarea acestei sume nu depinde de modul în care se face descompunerea poligonului în triunghiuri (există o infinitate de descompuneri de acest fel). De obicei, matematica școlară trece ușor peste lucruri de acest fel, dar în cărțile mai amănunțite și mai riguroase (cum este, de exemplu, tratatul de geometrie elementară (—citat în capitolul III — al marelui matematician francez Jacques Hadamard) aceste probleme sînt cercetate cu toată atenția. Neplăcut în aceste situație nu este atît faptul că astfel de probleme nu sînt rezolvate în cadrul manualelor școlare (probabil că cel puțin la clase mai mici nici nu este cazul să se facă așa ceva), cît faptul că nu reușim să dezvoltăm la elevi (și uneori nici la studenți!) acel grad de exigență care să-i conducă să-și pună singuri asemenea probleme. Exigența față de definiții este foarte scăzută la cei mai mulți începători în matematică. Chiar mulți dintre elevii și studenții sîrguincioși, care urmăresc cu grijă desfășurarea unei demonstrații matematice, căutînd să-i înțeleagă fiecare pas, meditează insuficient asupra definițiilor care se introduc în matematică, hrănindu-se (poate subconștient) cu ideea greșită că definițiile matematice sînt arbitrare. În fapt, o definiție matematică sintetizează o îndelungată experiență intuitivă și logică, ea constituie una dintre modalitățile prin care matematica își trage seva din lumea înconjurătoare.

Să ne întoarcem acum la problema măsurii. Este limpede că punctul de plecare îl vor constitui mulțimile pentru care știm să definim măsura. Pe dreaptă, singurele mulțimi de acest fel sînt intervalele. Însă, de aici nu-i decît un pas pînă la definiția unei măsuri pentru orice mulțime a acestei drepte, care se poate reprezenta ca o reuniune finită sau numărabilă de intervale. O astfel de mulțime A se poate reprezenta totdeauna ca o reuniune finită sau numărabilă de intervale I_n disjuncte. Notînd cu a_n și b_n extremitățile intervalului I_n ($n = 1, 2, \ldots$), vom defini măsura $m(A)$ a mulțimii A prin egalitatea

$$m(A) = \sum_{n=1}^{\infty} (b_n - a_n). \tag{10}$$

Se poate arăta că valoarea $m(A)$ nu depinde de modul în care mulțimea A este descompusă în intervale disjuncte. Prin definiția pusă n-am făcut decît să postulăm, să impunem (într-un caz particular) proprietatea de aditivitate a măsurii, proprietate care își are rădăcinile în experiența intuitivă pe care au dobîndit-o oamenii relativ la operația de măsurare.

Fie G o mulțime a dreptei care se poate reprezenta ca o reuniune finită sau numărabilă de intervale deschise și disjuncte două cîte două. Despre G se spune că este o mulțime deschisă. Ei i se atribuie o măsură prin relația (10) de mai sus, deoarece orice mulțime deschisă este de tipul mulțimilor A la care ne-am referit.

Însă nu orice parte a dreptei numerice este o mulțime deschisă. De exemplu, mulțimea punctelor de abscisă rațională nu este deschisă. Vom încerca atunci un proces de aproximare a unei mulțimi arbitrare prin mulțimi deschise. Fie E o mulțime oarecare a dreptei numerice. Există cel puțin o mulțime deschisă care conține pe E: aceasta este dreapta numerică. Să considerăm, pentru fiecare mulțime deschisă G care conține pe E, măsura $m(G)$ definită de (10). Deoarece, oricare ar fi G, valoarea $m(G)$ este nenegativă, rezultă că mulțimea valorilor $m(G)$ obținute pentru toate mulțimile deschise G care conțin pe E este o mulțime mărginită inferior și deci, conform unei teoreme cunoscute din analiza matematică elementară, ea admite o margine inferioară. Această margine inferioară o vom nota $m_e(E)$ și o vom numi măsura exterioară a mulțimii E:

$$m_e(E) = \inf_{G \supseteq E} m\ (G). \tag{11}$$

Următoarele două proprietăți pot fi ușor verificate de către cititor: 1) dacă A este un interval, atunci $m(A)$ dată de (10) coincide cu lungimea obișnuită a lui A; 2) pentru orice mulțime A, care este o reuniune finită sau numărabilă de intervale disjuncte (în particular, pentru orice mulțime A deschisă), avem $m(A) = m_e(A)$.

Pentru a ajunge la conceptul de măsurabilitate Lebesgue, ne vom concentra atenția asupra diferenței $G—E$, unde E este o mulțime arbitrară, iar G este o mulțime deschisă

arbitrară, care conține pe E. Deoarece conceptul de măsură exterioară are sens pentru orice mulțime a dreptei numerice, el va avea sens și pentru $G—E$; deci are sens expresia $m_e(G—E)$. Măsurabilitatea mulțimii E revine, în concepția lui Lebesgue, la posibilitatea de a alege mulțimea deschisă G care conține pe E în așa fel, încît măsura exterioară a lui $G—E$ să fie oricît de mică dorim. Mai precis, mulțimea E este măsurabilă (în sensul lui Lebesgue), dacă pentru fiecare număr pozitiv ε există o mulțime deschisă G care conține pe E și pentru care $m_e(G—E) < \varepsilon$. Pentru orice mulțime E măsurabilă Lebesgue măsura ei Lebesgue este egală, prin definiție, cu măsura exterioară a lui E și se notează $m(E)$.

Un întreg program de cercetare se impune acum. Astfel, se demonstrează că orice interval și orice mulțime deschisă sînt mulțimi măsurabile Lebesgue. Măsura Lebesgue a unui interval coincide cu lungimea sa, iar măsura Lebesgue a unei mulțimi deschise coincide cu cea dată de (10). Fiind dat un șir $E_1, E_2, \ldots, E_n, \ldots$, de mulțimi măsurabile Lebesgue, reuniunea lor E este de asemenea măsurabilă Lebesgue. Dacă în plus, mulțimile E_n sînt disjuncte două cîte două, atunci $m(E) = m(E_1) + m(E_2) - \ldots + m(E_n) + \ldots$ (proprietatea de aditivitate numărabilă a măsurii Lebesgue). Dacă mulțimile A și B sînt măsurabile Lebesgue și dacă $A \subset B$, atunci $m(A) \leqslant m(B)$ (proprietatea de monotonie a măsurii Lebesgue); într-adevăr, avem $B = A \cup (B - A)$ deci, în baza proprietății de aditivitate, $m(B) = m(A) + M(B - A)$ și, ținînd seama că măsura Lebesgue nu poate lua valori negative, rezultă că $m(A) \leqslant m(B)$. Să mai observăm că orice mulțime A de măsură nulă (în sensul considerat în teorema lui Lebesgue de integrabilitate riemanniană) este măsurabilă Lebesgue și $m(A) = 0$. Reciproc, dacă $m(A) = 0$, atunci A este de măsură nulă.

CÎT DE LARGĂ ESTE CLASA MULȚIMILOR MĂSURABILE LEBESGUE?

Acum, după ce cunoaștem conceptul de măsurabilitate introdus de Lebesgue, se pune în mod natural următoarea întrebare: Este acest concept suficient de larg pentru ca

orice mulțime a dreptei numerice să fie măsurabilă? În cazul contrar, este clasa mulțimilor măsurabile Lebesgue suficient de largă, de bogată, pentru ca ea să înglobeze toate mulțimile care intervin în problemele importante de matematică, mulțimi a căror întindere se cere a fi măsurată?

Să ne ocupăm mai întîi de prima întrebare. Răspunsul la ea este negativ, după cum a arătat la începutul secolului nostru (nu mult timp după introducerea de către Lebesgue a conceptului de măsurabilitate) matematicianul italian Vitali. Demonstrația lui Vitali este atît de ingenioasă și de instructivă, atît de semnificativă din punct de vedere metodologic și chiar filozofic, încît nu ne putem opri de a o expune aici.

Înainte de a trece la demonstrația lui Vitali, vom menționa o proprietate foarte importantă a măsurii Lebesgue, proprietate care va interveni esențial în cele ce urmează:

Dacă mulțimea H se obține din mulțimea E printr-o translație și dacă E este măsurabilă Lebesgue, atunci H este de asemenea măsurabilă Lebesgue și $m(H) = m(E)$.

Într-adevăr, este suficient să observăm mai întîi că teorema este adevărată în cazul în care E și H sînt intervale. De aici rezultă că teorema este adevărată și pentru mulțimi deschise, deoarece măsura unei mulțimi deschise se exprimă, prin relația (10), cu ajutorul lungimilor unor intervale. În sfîrșit, din faptul că H se obține din E printr-o translație, rezultă că orice mulțime deschisă care conține pe H (sau pe E) se obține, printr-o translație, dintr-o mulțime deschisă care conține pe E (respectiv pe H), și teorema rezultă în toată generalitatea ei.

Vom numi proprietatea de mai sus *proprietatea de invarianță a măsurii Lebesgue în raport cu o translație.*

Să considerăm mulțimea A a punctelor din intervalul $[0, 1]$ și să introducem în A o relație de echivalență, după cum urmează. Două puncte din A sînt echivalente, dacă abscisele lor diferă printr-un număr rațional. Ținînd seamă că mulțimea numerelor raționale este numărabilă, rezultă că fiecare clasă de echivalență este o mulțime numărabilă. Însă mulțimea A este, după cum se știe, nenumărabilă; deci mulțimea claselor de echivalență este și ea nenumărabilă (deoarece o reuniune numărabilă de mulțimi numă-

rabile este de asemenea o mulțime numărabilă). Presupunînd că s-a introdus în fiecare clasă de echivalență o anumită corespondență biunivocă cu mulțimea numerelor naturale, să notăm cu A_n mulțimea care conține, din fiecare clasă de echivalență, exact acel element care este asociat numărului natural n. Rezultă că fiecare mulțime A_n este nenumărabilă, și

$$A = \bigcup_{n=1}^{\infty} A_n. \tag{12}$$

Vom demonstra că printre mulțimile A_n există cel puțin una care nu este măsurabilă Lebesgue. Vom proceda prin reducere la absurd. Să admitem că toate mulțimile A_n sînt măsurabile Lebesgue. Dacă am avea $m(A_n) = 0$ pentru $n = 1, 2, \ldots$ ar rezulta, în baza proprietății de aditivitate numărabilă a măsurii Lebesgue, că $m(A) = 0$, ceea ce ar veni în contradicție cu faptul că A este un interval de lungime egală cu 1. Există deci cel puțin o valoare a lui n — fie ea p (dar nu putem preciza valoarea lui p) — pentru care $m(A_p) > 0$. Să punem, pentru simplitate, $A_p = B$. Avem deci $m(B) > 0$. Să notăm cu B_n mulțimea obținută din B printr-o translație de mărime $\frac{1}{n}$. (Aceasta înseamnă că abscisele punctelor din B_n sînt de forma $x + \frac{1}{n}$, unde x parcurge pe B.) În baza proprietății de invarianță a măsurii Lebesgue în raport cu o translație, rezultă că fiecare mulțime B_n este măsurabilă Lebesgue și $m(B_n) = m(B)$ $(n = 1, 2, \ldots)$.

Vom arăta acum că mulțimile B_n sînt disjuncte două cîte două. Într-adevăr, să admitem că există un punct a comun mulțimilor B_s și B_t, deși $s \neq t$. Din apartenența lui a la B_s rezultă existența unui element b din B, pentru care $a = b + \frac{1}{s}$. Din apartenența lui a la B_t rezultă existența unui element c din B, pentru care $a = c + \frac{1}{t}$. Rezultă că $b + \frac{1}{s} = c + \frac{1}{t}$, unde s și t sînt numere

naturale, deci diferența $b - c$ este un număr rațional, deci b și c aparțin aceleiași clase de echivalență. Dar să nu uităm că B este egală cu mulțimea A_p, deci cu o mulțime care conține cîte un element *și numai cîte unul* din fiecare clasă de echivalență; deci $b = c$, de unde rezultă că $s = t$. Am ajuns astfel la o contradicție (prin ipoteză, $s \neq t$) deci mulțimile B_n sînt disjuncte două cîte două.

O ultimă observație relativă la mulțimea B_n: din faptul că B este conținută în A rezultă că fiecare mulțime B_n este conținută în intervalul [0, 2].

Folosind proprietățile stabilite mai sus și bazîndu-ne pe proprietatea de aditivitate numărabilă a măsurii Lebesgue, rezultă (notînd cu D reuniunea mulțimilor B_n): $m(D) = m(B_1) + m(B_2) + \ldots + m(B_n) + \ldots = m(B) + + m(B) + \ldots + m(B) + \ldots = + \infty$. Însă mulțimea D este conținută în intervalul [0,2], deci, în baza proprietății de monotonie a măsurii Lebesgue, măsura Lebesgue a mulțimii D nu poate fi mai mare ca 2. Am ajuns astfel la o contradicție, care arată că cel puțin una dintre mulțimile A_n (dar nu știm care anume dintre ele!) nu este măsurabilă Lebesgue. Teorema lui Vitali este astfel demonstrată.

Să facem acum cîteva observații pe marginea demonstrației de mai sus. Ceea ce atrage în primul rînd atenția este caracterul ei neefectiv, neconstructiv. Demonstrația nu furnizează un exemplu individual de mulțime nemăsurabilă Lebesgue. Trebuie să spunem că matematicianul polonez Waclaw Sierpiński, reluînd și aprofundînd demonstrația de mai sus, a reușit să arate, între altele, că toate mulțimile A_n sînt nemăsurabile Lebesgue. Dar nici prin acest fapt existența mulțimilor nemăsurabile Lebesgue nu a căpătat un caracter constructiv. Într-adevăr, să ne amintim cum am obținut mulțimile A_n. A fost necesar, în prealabil, să fixăm în fiecare clasă de echivalență o anumită corespondență biunivocă cu mulțimea numerelor naturale. Dar există o infinitate de clase de echivalență, deci nu putem lua pe rînd fiecare clasă pentru a alege, pentru fiecare în parte, o corespondență biunivocă cu mulțimea numerelor naturale, deoarece în felul acesta nu am termina niciodată. Totuși, noi am admis existența acestei infinități de alegeri, fără de care demonstrația nu ar fi putut continua. În postula-

rea existenței acestei infinități de alegeri se ascunde unul din faptele primare, deosebit de profunde, ale raționamentelor cu mulțimi infinite, fapt descoperit de către matematicianul german Ernst Zermelo în primul deceniu al secolului nostru. Zermelo a observat că numeroase modalități de raționament utilizate în studiul mulțimilor infinite revin, în ciuda aparentei lor diversități, la una și aceeași schemă:

Fiind dată o familie $\mathcal{F}$ de mulțimi nevide și disjuncte, există o mulțime A care conține cîte un element și numai unul din fiecare mulțime a lui $\mathcal{F}$.

Multă vreme matematicienii au încercat să demonstreze acest enunț, să-l transforme într-o teoremă, dar toate încercările au eșuat și enunțul a rămas în matematică sub denumirea de *Axioma alegerii* sau *axioma lui Zermelo*. Nu este greu de văzut că existența mulțimilor A_n din demonstrația lui Vitali se bazează tocmai pe utilizarea axiomei lui Zermelo.

Pentru a da o imagine mai clară despre natura acestei axiome, să considerăm următorul exemplu. Să asociem fiecărei funcții reale f, de o variabilă reală, funcția opusă $-f$. Să considerăm mulțimea M a cuplelor $<f, -f>$ și să ne propunem să construim o mulțime N care conține cîte un element și numai unul din fiecare cuplu. Încercarea de a *construi* o astfel de mulțime nu reușește, dar axioma alegerii ne asigură de *existența* mulțimii N.

Se pune în mod natural întrebarea dacă nu cumva, folosind o altă metodă de demonstrație decît cea propusă de Vitali, nu am putea stabili, pe o cale constructivă, existența mulțimilor nemăsurabile Lebesgue. Dar, deși cunoaștem în prezent numeroase căi care conduc la existența mulțimilor nemăsurabile Lebesgue, nici una din aceste căi nu poate evita folosirea axiomei alegerii. Cu alte cuvinte mulțimile nemăsurabile Lebesgue sînt dincolo de granițele matematicii efective, constructive.

Astăzi, natura logică a axiomei alegerii este complet elucidată. În urmă cu aproape trei decenii, matematicianul de origine austriacă Kurt Gödel a demonstrat că axioma alegerii nu poate fi infirmată, în sensul că dacă sistemul de axiome al teoriei mulțimilor este necontradictoriu (nici acest fapt nu este încă elucidat!), atunci el rămîne necon-

tradictoriu prin adăugarea, ca o nouă axiomă, a axiomei alegerii. În urma rezultatului lui Gödel, devenea clar că nu ne puteam aștepta la o teoremă care să afirme falsitatea axiomei alegerii. Mai rămînea să se lămurească situația negației axiomei alegerii. În urmă cu cîțiva ani, matematicianul american Paul Cohen a rezolvat această problemă, demonstrînd că dacă sistemul de axiome al teoriei mulțimilor este necontradictoriu, atunci el rămîne necontradictoriu prin adăugarea, ca o nouă axiomă, a negației axiomei alegerii. Pentru acest rezultat, precum și pentru alte rezultate, Paul Cohen a fost distins la Congresul internațional al matematicienior din 1966, de la Moscova, cu medalia Fields, cea mai înaltă distincție internațională care se acordă, pentru descoperiri în domeniul matematicii, exclusiv matematicienilor sub 40 de ani.

Combinînd rezultatele lui Gödel și Cohen, rezultă că axioma alegerii constituie o propoziție independentă, în sensul că nici ea, nici negația ei nu sînt contradictorii. Avem deci două matematici, una care adoptă axioma alegerii, alta care adoptă negația ei. Existența mulțimilor nemăsurabile Lebesgue a fost stabilită în cadrul primei.

Situația axiomei alegerii este deci asemănătoare, în teoria mulțimilor, cu aceea a postulatului lui Euclid în geometrie; rezultatele lui Gödel și Cohen corespund descoperirilor lui Lobacevski și Bolyai. Însă, în timp ce domeniul geometriilor neeuclidiene este destul de clar conturat, avem deocamdată o imagine destul de confuză despre cele două matematici corespunzătoare adoptării, respectiv neadoptării axiomei alegerii. În bună măsură, această situație este datorită faptului că nu dispunem de o evidență clară a propozițiilor matematice în a căror demonstrație intervine axioma alegerii și, chiar în cazul multor propoziții despre care sîntem conștienți că au fost stabilite cu ajutorul axiomei alegerii, nu știm dacă ele n-ar putea fi stabilite și fără această axiomă.

Teorema lui Vitali ne atrage atenția asupra dificultății de a ne întîlni cu mulțimi nemăsurabile Lebesgue. Această situație este confirmată de un ansamblu de teoreme relative la comportarea mulțimilor măsurabile Lebesgue față de diferite operații cu mulțimi. Astfel, se demonstrează că

orice reuniune finită sau numărabilă și orice intersecție finită sau numărabilă de mulțimi măsurabile Lebesgue sînt de asemenea mulțimi măsurabile Lebesgue, iar diferența a două mulțimi măsurabile Lebesgue este măsurabilă Lebesgue. Orice mulțime finită sau numărabilă este de măsură Lebesgue nulă, orice interval este o mulțime măsurabilă Lebesgue. Mulțimile întîlnite în mod curent în analiză se obțin, pornind de la intervale, prin aplicarea de un număr oarecare de ori a operațiilor de reuniune finită sau numărabilă, intersecție finită sau numărabilă și trecere la complementară. Astfel de mulțimi se numesc boreliene, după numele matematicianului francez Emile Borel, care le-a introdus. Borel a construit și o teorie a măsurii pentru mulțimile care-i poartă numele, teorie care ia ca punct de plecare măsura elementară a intervalelor. După cum s-a putut arăta, orice mulțime boreliană este măsurabilă Lebesgue, iar măsura ei Borel coincide cu măsura ei Lebesgue. Există însă mulțimi măsurabile Lebesgue care nu sînt boreliene. Nu insistăm asupra acestui fapt, deoarece el va fi tratat în capitolul V, unde se va arăta că unele probleme de analiză conduc la considerarea unor mulțimi măsurabile Lebesgue neboreliene. Se poate arăta că familia mulțimilor boreliene este de puterea continuului, în timp ce familia mulțimilor măsurabile Lebesgue este de putere superioară continuului; cu alte cuvinte, prima dintre aceste familii poate fi pusă în corespondență biunivocă cu mulțimea numerelor reale, în timp ce a doua dintre ele nu poate fi pusă în corespondență biunivocă cu nici o mulțime de numere reale.

CÎTE CEVA DESPRE MĂSURABILITATEA LEBESGUE A FUNCȚIILOR

Acum, după ce am dobîndit anumite cunoștințe despre măsurabilitatea și măsura Lebesgue a mulțimilor, să ne întoarcem la problema care a cerut studiul acestei măsurabilități, anume la problema măsurabilității funcțiilor, așa cum a fost pusă ea de către Lebesgue. Ne amintim că succesul metodei lui Lebesgue era condiționat de posibilitatea de a lărgi cît mai mult clasa mulțimilor pentru care

putem defini în mod natural o măsură, în așa fel încît cît mai multe dintre funcțiile uzuale să fie măsurabile.

Ne așteptăm acum ca măsura Lebesgue să dea un răspuns satisfăcător acestei probleme, deoarece mulțimile întîlnite în mod obișnuit în analiză sînt măsurabile Lebesgue. Aceasta face ca și funcțiile întîlnite în mod obișnuit să fie măsurabile Lebesgue, adică să fie de așa natură, încît mulțimile de forma $\{x: a \leqslant f(x) < b\}$ (aceste mulțimi se numesc *mulțimile Lebesgue asociate funcției f*) să fie măsurabile Lebesgue, oricare ar fi a și b (finite sau infinite). Să schițăm, pentru unele tipuri de funcții uzuale, modul în care se stabilește măsurabilitatea lor Lebesgue. Dacă f este continuă, orice mulțime Lebesgue a ei este intersecția dintre o mulțime închisă și una deschisă, deci este măsurabilă Lebesgue. Dacă f este monotonă, mulțimile ei Lebesgue sînt intervale, deci mulțimi măsurabile Lebesgue. Se poate arăta că operațiile algebrice de sumă, diferență și produs, aplicate funcțiilor măsurabile Lebesgue, conduc tot la funcții măsurabile Lebesgue. Raportul dintre două funcții măsurabile Lebesgue, dintre care a doua nu se anulează, este de asemenea o funcție măsurabilă Lebesgue. Aceste proprietăți sînt naturale, ele nu constituie o surpriză, deoarece le-am întîlnit și la funcții continue, și la funcții derivabile, și la alte clase de funcții studiate în analiză. Există însă o proprietate care distinge net măsurabilitatea Lebesgue de proprietățile de continuitate, derivabilitate sau integrabilitate riemanniană. Se știe că nici una dintre aceste din urmă proprietăți nu este invariantă prin trecere la limită; limita unui șir convergent de funcții continue (sau derivabile, sau integrabile Riemann) nu este obligatoriu o funcție continuă (respectiv derivabilă, integrabilă Riemann). Se poate arăta însă că limita unui șir convergent de funcții măsurabile Lebesgue este totdeauna o funcție măsurabilă Lebesgue. Acest fapt conferă o deosebită simplitate și eleganță clasei funcțiilor măsurabile Lebesgue; această clasă este închisă nu numai față de operațiile algebrei, dar și față de operațiile analizei (trecerea la margine inferioară sau superioară; trecerea la limită inferioară sau superioară). Dintre operațiile uzuale, numai una ne poate

scoate din clasa funcțiilor măsurabile Lebesgue: operația de suprapunere. Din faptul că f este măsurabilă Lebesgue pe $[a, b]$, iar g este măsurabilă Lebesgue pe un interval care conține valorile lui f, nu rezultă că funcția h dată de $h(x) = g(f(x))$ este măsurabilă Lebesgue pe $[a, b]$. Mai mult decît atît, se poate arăta că orice funcție reală de o variabilă reală este o suprapunere a două funcții măsurabile Lebesgue. Acest fapt produce dificultăți în stabilirea teoremelor de schimbare de variabilă la integrala Lebesgue (se știe că în aceste teoreme intervine operația de suprapunere), dar dificultăți similare apar și atunci cînd vrem să stabilim teoreme cît mai generale de schimbare de variabilă la integrala Riemann, deoarece nici suprapunerea a două funcții integrabile Riemann nu este totdeauna o funcție integrabilă Riemann (mai mult decît atît, se poate arăta că orice funcție reală, mărginită, de o variabilă reală, este suprapunerea a două funcții integrabile Riemann). Deci nu putem interpreta această dificultate ca o inferioritate a integralei Lebesgue față de integrala Riemann.

Folosind invarianța măsurabilității Lebesgue față de operațiile algebrei și analizei, putem stabili noi clase de funcții uzuale măsurabile Lebesgue. Astfel, orice funcție cu variație mărginită este măsurabilă Lebesgue, fiind diferența a două funcții monotone (deci măsurabile Lebesgue). Orice funcție derivată este măsurabilă Lebesgue, fiind limita unui șir convergent de funcții continue (deci măsurabile Lebesgue). Orice funcție semicontinuă este măsurabilă Lebesgue, fiind limita unui șir de funcții continue. Se poate arăta că și funcțiile integrabile Riemann sînt măsurabile Lebesgue. Deoarece, după cum am văzut, orice funcție mărginită, măsurabilă Lebesgue, este integrabilă Lebesgue și, ținînd seama că orice funcție integrabilă Riemann este mărginită, rezultă că orice funcție integrabilă Riemann este integrabilă Lebesgue. Se poate arăta că integrala Riemann este egală cu integrala Lebesgue a funcției considerate, ceea ce arată că integrala Lebesgue este o extensiune a integralei Riemann. Această extensiune este efectivă, deoarece există funcții integrabile Lebesgue care nu sînt integrabile Riemann. Astfel, funcția egală cu 1

pentru x raţional şi egală cu zero pentru x iraţional nu este integrabilă Riemann, pe nici un interval, deoarece ea e discontinuă în orice punct, dar este integrabilă Lebesgue pe orice interval, deoarece mulţimea numerelor raţionale şi mulţimea numerelor iraţionale sînt mulţimi boreliene, deci măsurabile Lebesgue şi, după cum se poate vedea uşor, orice mulţime Lebesgue a funcţiei considerate se obţine din aceste două mulţimi prin operaţii algebrice simple.

Trebuie să observăm totuşi că chiar unele proprietăţi studiate în analiza clasică pot conduce la funcţii nemăsurabile Lebesgue. Nu vom insista aici asupra acestui fapt, menţionăm doar că există funcţii cu proprietatea lui Darboux care nu sînt măsurabile Lebesgue pe nici un interval.

Modul în care s-a definit măsurabilitatea Lebesgue a unei funcţii nu lasă să se vadă uşor în ce constă structura intimă a unei funcţii măsurabile Lebesgue. Această structură n-a putut fi detectată decît cu mari eforturi, concretizate într-o teoremă a lui Luzin şi o alta datorită lui Denjoy şi Stepanov. Teorema lui Luzin afirmă, în esenţă, că orice funcţie măsurabilă Lebesgue admite o restricţie continuă pe o mulţime de măsură Lebesgue oricît de apropiată de aceea a intervalului de definiţie a funcţiei. (Fiind dată o funcţie $f: A \to B$ şi o mulţime $D \subset A$, prin restricţia lui f la D se înţelege funcţia $\varphi: D \to B$, pentru care $\varphi(x) = f(x)$, oricare ar fi $x \in D$. Mai precis, dacă f este măsurabilă Lebesgue pe $[a, b,]$ atunci pentru orice număr pozitiv ε există o mulţime închisă F conţinută în $[a, b]$, astfel încît restricţia lui f la F să fie continuă iar măsura Lebesgue a diferenţei $[a, b] - F$ să fie mai mică decît ε. Precizăm că prin continuitatea restricţiei lui f la F se înţelege faptul că pentru orice x din F şi pentru orice şir x_n de puncte din F care tinde către x şirul $f(x_n)$ tinde către $f(x)$.

Teorema lui Luzin permite să se compare din punct de vedere structural funcţiile integrabile Riemann şi cele măsurabile Lebesgue. În timp ce primele sînt continue aproape peste tot, celelalte au o proprietate mai slabă, în sensul că proprietatea de continuitate nu are loc obligatoriu aproape peste tot, ci doar pe o mulţime de măsură oricît de apropiată de aceea a intervalului de definiţie; în plus, la funcţii măsurabile continuitatea nu se referă la funcţia

considerată, ci la restricția ei la o anumită mulțime închisă.

Putem micșora deșeul din teorema lui Luzin, acea mulțime $[a, b] - F$ de măsură inferioară lui ε. Însă aceasta nu se poate obține decît cu prețul unei slăbiri a proprietății de continuitate, anume înlocuind continuitatea obișnuită prin continuitatea aproximativă sau asimptotică (definiția acestei noțiuni a fost dată în capitolul I). După cum s-a menționat în capitolul I, Denjoy a demonstrat că orice funcție măsurabilă este aproximativ continuă aproape peste tot. Stepanov a demonstrat că și reciproca este adevărată: orice funcție aproximativ continuă aproape peste tot este măsurabilă Lebesgue. Deci măsurabilitatea Lebesgue a unei funcții este echivalentă cu continuitatea ei aproximativă aproape peste tot.

O COMPARAȚIE ÎNTRE INTEGRALA RIEMANN ȘI INTEGRALA LEBESGUE

Am văzut în paragraful precedent că integrala Lebesgue constituie o veritabilă extensiune a integralei Riemann. Totuși, din prezentarea dată pînă aici nu putem căpăta încă o imagine mai clară a imensului progres pe care l-a realizat Lebesgue în teoria integralei. Acest progres este atît de mare, încît chiar în multe probleme relative la integrala Riemann metodele promovate de Lebesgue duc mai repede și mai elegant la găsirea soluției.

În primul rînd, trebuie să eliminăm două restricții adoptate în expunerea de mai sus. „Teritoriul" natural pe care se studiază o funcție măsurabilă Lebesgue (și deci și integrala Lebesgue) nu este un interval, ci o mulțime măsurabilă Lebesgue. O examinare atentă a definițiilor date conceptelor de funcție măsurabilă Lebesgue și de integrală Lebesgue arată că nicăieri nu a intervenit faptul că teritoriul pe care se lucra era presupus un interval, ci doar proprietatea intervalelor de a fi mulțimi măsurabile Lebesgue.

Nici restricția mărginirii funcției nu este esențială. Se poate extinde conceptul de integrabilitate Lebesgue

și la funcții măsurabile nemărginite. Fie mai întîi o funcție f măsurabilă Lebesgue și pozitivă pe mulțimea măsurabilă E. Să notăm cu f_n funcția egală cu f în punctele în care valoarea lui f este mai mică decît n și egală cu n în celelalte puncte din E. Șirul f_n este monoton crescător (datorită faptului că funcția f este pozitivă), iar funcțiile f_n sînt mărginite și măsurabile Lebesgue. Deoarece — după cum se poate arăta ușor — integrala Lebesgue posedă, ca și integrala Riemann, proprietatea de monotonie, rezultă că șirul care are ca termen de rang n integrala Lebesgue a funcției f_n pe mulțimea E este un șir monoton crescător, deci are limită. Valoarea acestei limite devine, prin definiție, integrala Lebesgue a funcției f pe mulțimea E. Dacă această valoare este finită, spunem despre funcția f că este integrabilă Lebesgue pe E.

Fie acum o funcție h măsurabilă Lebesgue, dar în rest arbitrară, definită pe mulțimea măsurabilă Lebesgue E. Ca orice funcție reală, și funcția h se poate scrie ca diferență a două funcții pozitive: $h = f - g$, unde $f = \frac{1}{2}(|h| + h)$ și $g = \frac{1}{2}(|h| - h)$. Nu este greu de văzut că funcțiile f și g sînt măsurabile Lebesgue. (Este suficient să observăm că modulul unei funcții măsurabile Lebesgue este de asemenea o funcție măsurabilă Lebesgue.) Vom spune că funcția h este integrabilă Lebesgue pe E, dacă funcțiile f și g sînt integrabile Lebesgue pe E. În acest caz, integrala Lebesgue a funcției h pe E va fi, prin definiție, egală cu diferența integralelor Lebesgue ale funcțiilor f și g pe E. Dacă funcția f (sau g) nu este integrabilă Lebesgue pe E, dar g (respectiv f) este integrabilă Lebesgue pe E, vom spune că integrala Lebesgue a funcției h este egală cu $+\infty$ (respectiv cu $-\infty$).

Despre o funcție integrabilă Lebesgue pe E în sensul generalizat de mai sus se obișnuiește uneori să se spună că este sumabilă pe E. O teoremă importantă din teoria integralei Lebesgue afirmă că o funcție măsurabilă Lebesgue pe E este sumabilă pe E, dacă și numai dacă modulul ei este o funcție sumabilă pe E.

După cum se știe, și integrala Riemann admite o extensiune la funcții nemărginite sau pe intervale nemărginite.

Numind integrabilitatea Riemann de acest tip „integrabilitate Riemann în sens generalizat", este natural să ne întrebăm cum stă acest tip de integrabilitate față de integrabilitatea Lebesgue, în forma extinsă din paragraful de față. Poate contrar așteptărilor, trebuie să spunem că integrabilitatea Riemann în sens generalizat *nu* implică sumabilitatea Lebesgue. Pentru a ilustra această aserțiune, să considerăm funcția

$$f(x) = \begin{cases} \frac{\sin x}{x}, & \text{dacă } x \neq 0 \\ 1, & \text{dacă } x = 0 \end{cases}.$$

Se știe că această funcție este integrabilă Riemann în sens generalizat pe $[0, \infty)$, dar modulul ei nu mai are această proprietate; mai precis, integrala Riemann în sens generalizat a funcției $|f|$ pe $[0, \infty]$ este egală cu $+\infty$. De aici rezultă, cum e ușor de văzut, că și integrala Lebesgue a funcției $|f|$ este egală cu $+\infty$ pe intervalul $[0, \infty)$, deci că funcția $|f|$ nu este sumabilă pe intervalul $[0, \infty)$. Însă, în baza teoremei amintite mai sus, aceasta implică nesumabilitatea Lebesgue a funcției f pe $[0, \infty)$.

Cu acest prilej, ne-am întîlnit cu o situație care marchează o deosebire esențială între integrala Riemann în sens generalizat și sumabilitatea Lebesgue. Noțiunea de semiconvergență, atît de importantă în studiul integralei Riemann în sens generalizat (a se vedea chiar exemplul funcției f de mai sus), este absentă din teoria integralei Lebesgue. Trebuie însă să remarcăm că chiar în teoria integralei Riemann în sens generalizat semiconvergența apare numai în studiul funcțiilor de o variabilă; în ceea ce privește funcțiile de două sau mai multe variabile, datorită caracterului mult mai restrictiv pe care-l dobîndește în acest caz noțiunea de convergență a unei integrale, o funcție f este integrabilă Riemann în sens generalizat pe un domeniu D, dacă și numai dacă $|f|$ este integrabilă Riemann în sens generalizat pe D.

Așa acum integrala Riemann poate fi studiată ca o funcție de intervalul pe care ea este definită, integrala Lebesgue poate fi și ea studiată ca o funcție de mulțimea mă-

surabilă pe care ea este definită. Se știe, în acest sens, că integrala Riemann este o funcție aditivă de interval. În ceea ce privește integrala Lebesgue, ea este aditivă nu numai ca funcție de interval, ci și ca funcție de mulțime măsurabilă Lebesgue și nu numai în raport cu un număr finit de mulțimi, ci și în raport cu o infinitate numărabilă de mulțimi măsurabile Lebesgue. Mai precis, dacă E_1, E_2, ..., E_n,... sînt mulțimi măsurabile Lebesgue disjuncte două cîte două, atunci, notînd cu E reuniunea lor, orice funcție f sumabilă pe E este sumabilă pe fiecare E_n și avem

$$\int_E f = \int_{E_1} f + \int_{E_2} f + \ldots + \int_{E_n} f + \ldots,$$

unde integrala este concepută în sensul lui Lebesgue.

Această proprietate de aditivitate imprimă o deosebită eleganță calculului unei integrale Lebesgue. Vom ilustra această aserțiune prin două exemple: Fie mai întîi funcția lui Dirichlet, la care ne-am mai referit în capitolul de față: $f(x) = 1$, dacă x este rațional, $f(x) = 0$, dacă x este irațional. Am văzut că această funcție este măsurabilă Lebesgue; fiind și mărginită, ea este integrabilă Lebesgue pe orice interval. Ne propunem să calculăm valoarea integralei lui f pe $[a, b]$. Fie A mulțimea punctelor de abscisă rațională din $[a, b]$ și fie B mulțimea punctelor de abscisă irațională din același interval. Mulțimile A și B sînt măsurabile Lebesgue și avem, în baza proprietății de aditivitate,

$$\int_a^b f = \int_A f + \int_B f.$$

Însă f este constant egală cu 1 pe mulțimea A și constant egală cu 0 pe B, deci

$$\int_A f = 0, \quad \int_B f = 0, \int_a^b f = 0.$$

Vom da acum un alt exemplu, care va arăta eficacitatea teoremei de aditivitate a integralei Lebesgue ca funcție de mulțime, chiar în cazul integralei Riemann. Fie funcția

g — deja considerată — egală cu 0 în origine și în orice punct irațional și egală cu $\frac{1}{q}$ în orice punct de abscisă $\frac{p}{q}$, unde p și q sînt numere întregi prime între ele, iar q este strict pozitiv. Funcția g este continuă în orice punct irațional; fiind și mărginită, ea este integrabilă Riemann pe orice interval compact al dreptei numerice. Desigur că am putea evalua integrala Riemann a funcției g pe un interval $[a, b]$ pornind chiar de la definiția integralei Riemann, adică formînd sumele Riemann corespunzătoare și studiind comportarea lor. Dar mai rapid și mai elegant se obține valoarea integralei Riemann a funcției g pe $[a, b]$, dacă privim această integrală în accepția ei de integrală Lebesgue. Într-adevăr, în acest caz, notînd cu E mulțimea punctelor de abscisă irațională din $[a, b]$, avem, în baza teoremei de aditivitate,

$$\int_a^b g = \int_E g + \int_{[a, b] - E} g .$$

Dar mulțimea $[a, b] - E$ este de măsură nulă și este ușor de văzut că pe o mulțime de măsură nulă integrala Lebesgue a oricărei funcții este egală cu zero. Pe de altă parte, prima integrală din membrul al doilea este și ea egală cu zero, deoarece funcția g este identic nulă pe E. Așadar, integrala lui g pe $[a, b]$ este egală cu zero.

O altă proprietate importantă a integralei Lebesgue ca funcție de mulțime este proprietatea de continuitate absolută. Mai întîi trebuie să observăm că integrabilitatea Lebesgue este o proprietate ereditară, în sensul că dacă ea are loc pe o anumită mulțime măsurabilă, atunci are loc pe orice parte măsurabilă a acesteia. Fiind dată o funcție $F(E)$ definită pentru părțile măsurabile ale lui A, vom spune că ea este absolut continuă pe A, dacă valoarea ei în modul devine oricît de mică dorim atunci cînd măsura submulțimii E a lui A este suficient de mică; mai precis, fiecărui număr pozitiv ε îi corespunde un număr pozitiv η, astfel încît din $E \subset A$ și $m(E) < \eta$ să rezulte $|F(E)| < \varepsilon$. O teoremă fundamentală afirmă că dacă

f este o funcție sumabilă pe mulțimea măsurabilă A, atunci integrala Lebesgue

$$F(E) = \int_E f$$

este o funcție de mulțime absolut continuă pe A. Această proprietate nu constituie o surpriză, deoarece o întîlnim, într-o altă variantă, și la integrala Riemann. Dacă funcția reală f de o variabilă reală este integrabilă — Riemann sau Lebesgue — pe $[a, b]$, atunci funcția

$$F(x) = \int_a^x f(t)\,dt$$

este absolut continuă pe $[a, b]$, în sensul că pentru orice $\varepsilon > 0$ există un număr $\eta > 0$, cu proprietatea că dacă $(a_1, b_1), (a_2, b_2), \ldots, (a_n, b_n)$ constituie un sistem finit de intervale disjuncte din $[a, b]$, pentru care $\sum_{i=1}^{n} (b_i - a_i) < \eta$, atunci $\sum_{i=1}^{n} (f(b_i) - f(a_i)) < \varepsilon$ (a se vedea și capitolul III).

Există însă o deosebire esențială între integrala Riemann și integrala Lebesgue, în favoarea acesteia din urmă. Se poate demonstra că aditivitatea numărabilă, împreună cu continuitatea absolută caracterizează funcțiile de mulțime care sînt integrale Lebesgue; mai precis, fiind dată o funcție $F(E)$ definită pentru toate părțile măsurabile ale mulțimii măsurabile A, numărabil aditivă și absolut continuă pe A, există o funcție reală f definită și sumabilă pe A, astfel încît

$$F(E) = \int_E f$$

pentru orice parte măsurabilă E a lui A. În schimb, există funcții absolut continue de punct, care nu sînt integrale Riemann și nu se cunoaște nici o caracterizare structurală a funcțiilor de interval care sînt integrale Riemann (chiar în sens generalizat).

O altă problemă în care se manifestă superioritatea integralei Lebesgue față de integrala Riemann este aceea

a trecerii la limită sub semnul integralei. Fiind dat un șir de funcții reale $\{f_n\}$ integrabile Riemann pe $[a, b]$, convergent pe $[a, b]$ către o funcție f, este posibil ca funcția f să nu fie integrabilă Riemann, iar în cazul în care f este integrabilă Riemann pe $[a, b]$, este posibil ca integrala funcției f pe $[a, b]$ să nu fie egală cu limita integralelor funcțiilor f_n ($n = 1, 2, \ldots$) pe $[a, b]$. Pentru ca aceste anomalii să nu se producă, este suficient ca în convergența șirului $\{f_n\}$ să se manifeste proprietatea de uniformitate, cu alte cuvinte pragul N din definiția convergenței acestui șir să depindă numai de ε, *nu* și de punctul x ales în $[a, b]$. Este adevărat că această condiție de uniformitate este doar suficientă — nu și necesară — pentru ca anomaliile semnalate să nu se producă, totuși nu se cunoaște nici o condiție suficientă mai simplă și totodată mai generală decît cea menționată. În schimb, o astfel de condiție mai simplă și mai generală pentru trecerea la limită sub semnul integralei a putut fi găsită pentru integrala Lebesgue, și ea se enunță în modul următor: dacă șirul de funcții f_n sumabile pe E converge pe E către funcția f și dacă funcțiile f_n sînt egal majorate în modul de către o funcție sumabilă g, atunci integrala funcției f pe E este egală cu limita integralelor funcțiilor f_n pe E.

În cazul particular în care E este intervalul $[a, b]$, iar șirul $\{f_n\}$ converge uniform pe $[a, b]$, putem găsi cu ușurință o funcție sumabilă g care majorează pe $[a, b]$ fiecare funcție $|f_n|$. Într-adevăr, fie $\varepsilon > 0$. În virtutea convergenței uniforme, există un număr natural N, astfel încît $|f_n(x) - f(x)| < \varepsilon$ pentru orice $n > N_\varepsilon$ și pentru orice x din $[a, b]$. Inegalitatea obținută se mai poate scrie $f(x) - \varepsilon < f_n(x) < f(x) - \varepsilon$. Să notăm cu $g(x)$ pe cel mai mare dintre numerele $|f_1(x)|$, $|f_2(x)|, \ldots, |f_{N_\varepsilon}(x)|$, $|f(x) - \varepsilon|$, $|f(x) - \varepsilon|$. Funcția g este sumabilă pe $[a, b]$ și avem, pentru orice numar natural n și pentru orice x din $[a, b]$, inegalitatea $|f_n(x)| \leqslant g(x)$. Am demonstrat astfel că dacă un șir $\{f_n\}$ de funcții sumabile pe $[a, b]$ converge uniform pe $[a, b]$ către funcția f, atunci integrala funcției f pe $[a, b]$ este tocmai limita integralelor funcțiilor f_n pe $[a, b]$. Există însa șiruri de funcții sumabile, convergente pe $[a, b]$, egal majorate în modul de o funcție su-

mabilă, dar care nu converg uniform pe $[a, b]$. Rezultă că pentru trecerea la limită sub semnul integralei Lebesgue dispunem de condiții practice mai puțin restrictive decît cele folosite în cazul integralei Riemann.

O altă problemă în care se manifestă superioritatea integralei Lebesgue față de integrala Riemann este aceea a raportului dintre derivată și integrabilitate. Ne amintim că familia derivatelor integrabile Riemann este efectiv mai largă decît familia derivatelor continue, deci în orice caz integrala Riemann constituie, în această problemă, un progres în raport cu etapa Cauchy. Dar, după cum am văzut, există derivate mărginite care nu sînt integrabile Riemann. Deoarece orice derivată este o funcție măsurabilă Lebesgue și orice funcție mărginită și măsurabilă Lebesgue este integrabilă Lebesgue, rezultă că orice derivată mărginită este integrabilă Lebesgue. Deci clasa derivatelor integrabile Lebesgue este efectiv mai largă decît clasa derivatelor integrabile Riemann.

Fie acum o funcție reală f integrabilă Riemann pe $[a, b]$. Să punem

$$F(x) = \int_a^x f(t)\,\mathrm{d}\,t.$$

O teoremă elementară din teoria integralei Riemann afirmă că F este derivabilă în orice punct x în care f este continuă și avem $F'(x) = f(x)$. Însă funcția f, fiind integrabilă Riemann, este continuă aproape peste tot, deci avem, aproape peste tot, $F'(x) = f(x)$. Constatăm deci că divorțul dintre primitivă și integrală, care apărea în teoria lui Riemann, nu este chiar atît de grav: integrala, încetînd de a fi o „primitivă peste tot", rămîne totuși o „primitivă aproape peste tot" a funcției care figurează sub semnul integralei.

Să vedem ce se întîmplă dacă funcția f este integrabilă Lebesgue pe $[a, b]$. În acest caz, funcția F este absolut continuă, și o teoremă clasică, datorită de asemenea lui Lebesgue, ne asigură că o astfel de funcție este derivabilă aproape peste tot. Mai mult decît atît, se arată că avem, aproape peste tot pe $[a, b]$, $F'(x) = f(x)$. După cum vedem, am obținut pe o cale diferită (această cale putea fi folosită și pentru integrala Riemann) același rezultat

ca și în cazul în care f este integrabilă Riemann pe $[a, b]$. Și aici, integrala este o „primitivă aproape peste tot" a funcției care se integrează. Dar există alte aspecte sub care, și aici, integrala Lebesgue își manifestă superioritatea față de integrala Riemann. Mai întîi, trebuie să observăm că mulțimile de măsură nulă sînt „de efect nul" în procesul de integrare Lebesgue, în sensul că schimbarea comportării unei funcții pe o mulțime de măsură nulă nu influențează nici măsurabilitatea funcției, nici integrabilitatea ei și nici valoarea integralei Lebesgue a funcției. În particular, este permis ca funcția să nici nu fie măcar definită (sau să fie definită în mod arbitrar) în unele puncte ale teritoriului de integrare, cu condiția ca aceste puncte să formeze o mulțime de măsură nulă. Acest fapt este deosebit de important, deoarece, pentru unele funcții des întîlnite în analiză, este asigurată existența derivatei aproape peste tot, dar nu peste tot. Așa se întîmplă în clasa funcțiilor cu variație mărginită, care conține subclase importante, cum ar fi: clasa funcțiilor monotone, clasa funcțiilor lipschitziene, clasa funcțiilor absolut continue, clasa funcțiilor singulare. (O funcție f este lipschitziană pe $[a, b]$, dacă există o constantă λ pentru care $|f(x) - f(y)| \leqslant \lambda |x - y|$, oricare ar fi $x, y \in [a, b]$. O funcție φ este singulară pe $[a, b]$, dacă este cu variație mărginită, iar derivata se anulează aproape peste tot, fără a fi identic nulă. Funcțiile absolut continue au fost definite în capitolul III). O teoremă clasică a lui Lebesgue afirmă că orice funcție f, cu variație mărginită pe $[a, b]$, este derivabilă aproape peste tot pe $[a, b]$. Este lipsit de sens să punem problema integrabilității Riemann a derivatei f' pe $[a, b]$, deoarece integrabilitatea Riemann pe $[a, b]$ cere ca funcția considerată să fie definită peste tot pe $[a, b]$. Este adevărat că am putea completa pe f' în punctele în care ea nu există, atribuindu-i valori arbitrare, de exemplu, considerînd funcția g egală cu f' acolo unde f' este definită și egală cu zero în rest. Dar nimic nu ne asigură că printr-o altă completare a funcției f', de exemplu, definind funcția h egală cu f' acolo unde f' există, și egală cu 1 în rest, funcțiile g și h împărtășesc aceeași soartă din punctul de vedere al inte-

grabilității Riemann; se poate întîmpla ca una dintre aceste funcții să fie integrabilă Riemann pe $[a, b]$, iar cealaltă nu. În schimb funcțiile g și h (și orice alte funcții obținute prin prelungirea lui f') împărtășesc aceeași soartă din punctul de vedere al integrabilității Lebesgue; ele sînt toate sumabile, atîta vreme cît f — așa cum am presupus — este cu variație mărginită pe $[a, b]$. Deci oricare ar fi funcția f cu variație mărginită pe $[a, b]$, are sens și este finită integrala Lebesgue a derivatei f' pe $[a, b]$, deși este posibil ca f' să nu fie definită în toate punctele din $[a, b]$.

Să mai observăm faptul că în timp ce integrala Riemann se definește numai pe un interval, integrala Lebesgue se poate defini pe orice mulțime măsurabilă A, iar funcția $F(E)$ (unde E este măsurabilă și conținută în A), dată de valoarea integralei Lebesgue pe mulțimea E, este o funcție de mulțime derivabilă aproape peste tot pe A, derivata F' (definită ca fiind egală cu $\lim F(E)/m(E)$ cînd diam $E \to 0$) fiind egală aproape peste tot chiar cu funcția care se integrează.

Proprietatea funcțiilor integrabile Lebesgue de a admite o „primitivă aproape peste tot" se menține și la unele clase de funcții măsurabile care nu sînt integrabile Lebesgue. Astfel, Luzin a demonstrat că fiind dată o funcție reală f, măsurabilă Lebesgue, definită (eventual numai aproape peste tot) pe un interval (a, b) al dreptei numerice, există o funcție F continuă pe (a, b), derivabilă aproape peste tot pe (a, b) și astfel încît $F'(x) = f(x)$ aproape peste tot pe (a, b). Cu alte cuvinte, teorema care afirmă că o funcție continuă admite primitivă își are corespondentul ei la funcții măsurabile, cu condiția de a se înlocui primitiva obișnuită prin „primitiva aproape peste tot". Situația aceasta apare oarecum firească, dacă ne amintim că o funcție reală, măsurabilă, de o variabilă reală, este aproximativ continuă aproape peste tot. Punctele de derivabilitate ale funcției F evocă punctele de continuitate aproximativă ale funcției f și invers; dar fenomenul are un caracter statistic, neputîndu-se pune semnul egalității între mulțimea punctelor de derivabilitate ale lui F și mulțimea punctelor de continuitate aproximativă ale lui f. Acest caracter statistic îl prezintă și legătura din-

tre o funcție f integrabilă Riemann pe $[a, b]$ și integrala Riemann $F(x)$ a lui f de la a la x. Se știe, într-adevăr, că F este derivabilă în orice punct de continuitate al lui f, dar nu se poate afirma că f este continuă în orice punct de derivabilitate al lui F.

Am văzut, într-un paragraf anterior al acestui capitol, că ceea ce a transformat integrala într-un instrument de calcul deosebit de perfecționat și eficace a fost descoperirea — de către Leibniz și Newton — a posibilității de a reduce calculul integralelor la calculul primitivelor. Această posibilitate, concretizată în celebra formulă a lui Leibniz și Newton, presupune însă atît existența integralei, cît și existența primitivei funcției considerate. Condițiile acestea sînt nu numai necesare, dar și suficiente pentru validitatea formulei lui Leibniz și Newton, atîta vreme cît integrala este concepută în sensul lui Riemann, iar primitiva este considerată în sensul obișnuit. Într-adevăr, am văzut că dacă f este o funcție derivabilă, cu derivata integrabilă Riemann pe $[a, b]$, atunci integrala lui f' de la a la b este tocmai diferența $f(b) - f(a)$. Se pune în mod natural întrebarea dacă formula lui Leibniz-Newton își menține validitatea atunci cînd primitiva în sensul obișnuit este înlocuită cu „primitiva aproape peste tot". Această problemă se pune atît pentru integrala Riemann, cît și pentru integrala Lebesgue. Mai întîi însă trebuie să observăm că în timp ce primitiva obișnuită este unic determinată — în afara unei constante aditive —, iar diferența $f(b) - f(a)$ nu depinde de valoarea acestei constante, deci nici de alegerea primitivei, cu totul altfel se prezintă situația în ceea ce privește „primitiva aproape peste tot". Fără să intrăm în detalii, vom semnala faptul că există funcții continue, monoton crescătoare și neconstante pe $[a, b]$, a căror derivată este egală cu zero aproape peste tot pe $[a, b]$ (amintim că existența aproape peste tot a derivatei este asigurată pentru orice funcție cu variație mărginită, deci, în particular, și pentru funcțiile monotone). O astfel de funcție se numește o funcție de tip Cantor. De aici rezultă imediat că dacă F este o „primitivă aproape peste tot" a funcției f pe $[a, b]$, iar G este o funcție de tip Cantor pe $[a, b]$, atunci $F + G$

este de asemenea o „primitivă aproape peste tot" a lui f pe $[a, b]$. Mai mult decît atît, există o infinitate de funcții de tip Cantor pe $[a, b]$. Deci diferența $F(b) - F(a)$ nu este independentă de alegerea primitivei aproape peste tot și formula lui Leibniz-Newton își pierde valabilitatea. Astfel, dacă f este funcția identic nulă pe $[a, b]$, iar F este o funcție de tip Cantor pe $[a, b]$, F va fi o „primitivă aproape peste tot" a lui f pe $[a, b]$, dar în timp ce integrala (Riemann sau Lebesgue) a lui f pe $[a, b]$ este evident egală cu zero, diferența $F(b) - F(a)$ este strict pozitivă (deoarece F nu este constantă, fiind totuși monoton crescătoare pe $[a, b]$). Așadar *nu* avem egalitate între diferența $F(b) - F(a)$ și valoarea integralei Lebesgue a funcției f pe $[a, b]$. Observăm aici o nouă deosebire între integrala Riemann și integrala Lebesgue, dar de data aceasta, deosebirea este în favoarea celei dintîi. În timp ce formula Leibniz-Newton este adevărată pentru orice derivată integrabilă Riemann, această formulă își pierde valabilitatea atunci cînd este vorba de o derivată integrabilă Lebesgue.

Totuși, deficiența semnalată mai sus nu este gravă, deoarece în clasa „primitivelor aproape peste tot" ale unei funcții sumabile există o subclasă în raport cu care formula lui Leibniz-Newton își recapătă valabilitatea; este subclasa acelor „primitive aproape peste tot" care sînt absolut continue. O teoremă importantă ne asigură că dacă două funcții absolut continue au aproape peste tot aceeași derivată, atunci ele diferă printr-o constantă. Deci, dacă f este sumabilă pe $[a, b]$, iar F și G sînt două „primitive aproape peste tot" absolut continue ale funcției f, atunci $F(b) - F(a) = G(b) - G(a)$. În plus, se arată că diferența $F(b) - F(a)$ este egală tocmai cu valoarea integralei Lebesgue a funcției f pe $[a, b]$. De exemplu, dacă f este o funcție egală cu zero aproape peste tot pe $[a, b]$, subclasa „primitivelor aproape peste tot" absolut continue ale lui f este formată din toate funcțiile constante pe $[a, b]$ și avem (notînd cu F o funcție constantă pe $[a, b]$), $F(b) - F(a) =$ = integrala lui f pe $[a, b] = 0$.

Ținînd seamă de rolul important al mulțimilor de măsură nulă în teoria integralei Lebesgue, este natural să

ne întrebăm dacă nu cumva orice funcție integrabilă Lebesgue devine integrabilă Riemann prin modificarea convenabilă a valorilor ei pe o mulțime de măsură nulă. Răspunsul este negativ. Într-adevăr, considerînd pe intervalul $[a, b]$ o mulțime P perfectă, rară și de măsură pozitivă, funcția egală cu 1 în punctele din P și cu 0 în rest este integrabilă Lebesgue pe $[a, b]$, dar nu coincide aproape peste tot cu nici o funcție integrabilă Riemann, deoarece mulțimea punctelor de discontinuitate coincide cu P și, deci, este de măsură pozitivă.

Să observăm, în încheierea acestui paragraf, că un aspect important al superiorității integralei Lebesgue în raport cu integrala Riemann a fost prezentat în capitolul III, paragrafele 11 și 12.

REPARAȚII DUPĂ O INJUSTIȚIE

Cititorului atent al paragrafelor precedente nu i-a scăpat probabil faptul că în timp ce teoria integralei Lebesgue avea la bază o teorie corespunzătoare a măsurii mulțimilor și funcțiilor, teorie pe care o valorificam înlocuind diviziunile orizontale (adică ale intervalului de definiție) prin diviziuni verticale (adică ale intervalului valorilor), teoria integralei Riemann era lipsită de aceste două avantaje. Totuși, măsura Lebesgue a mulțimilor putea fi angajată și fără a recurge la partiții verticale; era suficient ca fiind dată o mulțime A măsurabilă Lebesgue, să considerăm o partiție a ei în mulțimi $A_1, A_2, \ldots, A_n$ măsurabile Lebesgue și să formăm, relativ la o astfel de partiție. sumele de tip Riemann

$$\sum_{i=1}^{n} f(\xi_i)\, m(A_i),$$

unde f este o funcție reală definită pe A, iar ξ_i aparține lui A_i pentru $i = 1, 2, \ldots, n$. Mai departe, am fi putut defini integrabilitatea și integrala, studiind comportarea sumelor de mai sus, exact după metoda lui Riemann. După cum se observă, această cale reține, dintre cele două mari idei ale lui Lebesgue (înlocuirea diviziunilor

orizontale prin diviziuni verticale și adoptarea, pe cale explicită, a unei teorii a măsurii mulțimilor), numai una dintre ele. Consecințele adoptării acestei căi au fost studiate și s-a constatat că cel puțin pentru anumite clase particulare de funcții, cum ar fi funcțiile reale de o variabilă reală, conceptul de integrabilitate care se obține este exact cel al lui Lebesgue, iar integrala obținută este chiar integrala lui Lebesgue. O teorie de acest fel este numită, de obicei, „o teorie de tip Riemann a integralei Lebesgue". Această teorie arată că operația de înlocuire a diviziunilor orizontale prin cele verticale nu este necesară pentru a se obține progresul scontat de Lebesgue, cel puțin în cazul în care avem în vedere numai funcții reale de o variabilă reală. Se poate însă arăta că această operație de înlocuire nu este nici suficientă pentru a se obține progresul realizat de integrala Lebesgue. Într-adevăr, s-a putut construi o „teorie de tip Lebesgue a integralei Riemann", arătîndu-se că simpla înlocuire a partițiilor orizontale prin cele verticale — fără adoptarea măsurii introduse de Lebesgue, ci a unei măsuri mai rudimentare — duce tot la un concept de integrabilitate echivalent cu cel al lui Riemann. La obținerea acestui rezultat, o contribuție însemnată a fost adusă de către matematicianul român acad. prof. Miron Nicolescu, mai întîi într-un articol publicat în urmă cu peste treizeci de ani (*Sur les fonctions mesurables* (*J*) în „Bulletin des sciences mathématiques", vol. 57, 1933, pp. 276—281), apoi în volumul al treilea al tratatului domniei sale de analiză matematică.

Vom schița, în cele ce urmează, acest punct de vedere și rezultatele mai importante la care el conduce.

Încă de la sfîrșitul secolului trecut, integrala Riemann a fost legată de o anumită concepție despre ideea de măsură, concepție datorită lui Jordan. Pentru o mulțime E liniară de puncte, se ajunge la noțiunea de măsurabilitate Jordan în felul următor. Fiecărui sistem finit de intervale disjuncte, conținute în E, i se asociază suma lungimilor intervalelor sale, iar marginea superioară a sumelor astfel obținute constituie, prin definiție, lungimea interioară a mulțimii E și se notează $l_i(E)$. Fiecărui sis-

tem finit de intervale disjuncte, care acoperă mulțimea E, i se asociază suma lungimilor intervalelor sistemului, iar marginea inferioară a sumelor astfel obținute constituie, prin definiție, lungimea exterioară a mulțimii E și se notează $l_e(E)$. Chiar din definițiile introduse rezultă că lungimea interioară a unei mulțimi nu poate întrece lungimea ei exterioară; dacă avem $l_i(E) = l_e(E)$, spunem că mulțimea E are lungime (sau este măsurabilă Jordan), iar valoarea comună a lungimilor interioară și exterioară se numește lungimea mulțimii E și se notează $l(E)$.

Pentru mulțimi în plan se procedează într-un mod analog. Se definește aria interioară $a_i(E)$ a unei mulțimi E ca fiind marginea superioară a ariilor domeniilor poligonale conținute în E. Se definește apoi aria exterioară $a_e(E)$ a mulțimii E ca fiind marginea inferioară a ariilor domeniilor poligonale care conțin mulțimea E. Avem totdeauna $a_i(E) \leqslant a_e(E)$. Dacă avem chiar egalitate între cele două arii, spunem că E are arie (sau este măsurabilă Jordan), iar valoarea comună a ariilor interioară și exterioară se numește măsura Jordan sau aria mulțimii E și se notează prin $a(E)$.

Dintre proprietățile măsurii și măsurabilității Jordan, amintim aici doar cîteva:

Măsura Jordan este o funcție nenegativă și monotonă de mulțime măsurabilă Jordan (dacă A și B sînt măsurabile Jordan, $A \subset B$, atunci măsura Jordan a lui A nu o întrece pe aceea a lui B).

O mulțime este măsurabilă Jordan, dacă și numai dacă frontiera ei este de măsură Jordan egală cu zero.

Complementara unei mulțimi măsurabile Jordan este de asemenea măsurabilă Jordan.

Orice reuniune finită și orice intersecție finită de mulțimi măsurabile Jordan sînt de asemenea mulțimi măsurabile Jordan. Dacă, în plus, mulțimile sînt disjuncte două cîte două, atunci măsura Jordan a reuniunii este egală cu suma măsurilor Jordan ale termenilor. Cu alte cuvinte, măsura Jordan este o funcție finit-aditivă de mulțime măsurabilă Jordan. Aditivitatea numărabilă are și ea loc cu condiția (nu totdeauna îndeplinită!) ca reuniunea

mulțimilor considerate să fie o mulțime măsurabilă Jordan.

Nu este greu de văzut că măsurabilitatea Jordan a unei mulțimi E implică măsurabilitatea ei Lebesgue și egalitatea dintre $m(E)$ și măsura Jordan a lui E.

Pe dreaptă, mulțimea lui Cantor are lungimea egală cu zero, dar mulțimea punctelor raționale din (0, 1) nu are lungime, deoarece lungimea ei interioară este nulă, iar cea exterioară este egală cu 1. Acest exemplu arată că conceptul de măsurabilitate introdus de Lebesgue este efectiv mai general decît cel introdus de Jordan. În orice caz, orice interval are lungime (în sensul lui Jordan), deci sîntem în progres față de concepția elementară asupra lungimii.

Fenomene mai interesante au loc în plan. De exemplu, se poate arăta că orice mulțime convexă din plan are arie. Pe dreaptă, teorema aceasta este adevărată în mod banal, deoarece singurele mulțimi convexe sînt aici intervalele. Tot în plan are loc o teoremă care dezvăluie legătura profundă dintre integrabilitatea Riemann și măsurabilitatea Jordan. Fiind dată o funcție reală f, definită și pozitivă pe intervalul $[a, b]$, să notăm cu D_f mulțimea punctelor ale căror coordonate x, y satisfac inegalitățile $a \leqslant x \leqslant b$ și $0 \leqslant y \leqslant f(x)$. O teoremă clasică, datînd încă de la sfîrșitul secolului trecut, afirmă că integrabilitatea Riemann a funcției f pe intervalul $[a, b]$ este echivalentă cu măsurabilitatea Jordan a mulțimii D_f. Dacă f este integrabilă Riemann pe $[a, b]$, integrala Riemann a lui f pe $[a, b]$ este egală cu aria mulțimii D_f.

Acest rezultat — ca și altele, pe care nu le mai semnalăm aici — a pus problema dacă nu cumva teoria integralei Riemann n-ar putea fi clădită pe teoria măsurii Jordan, după modelul folosit de Lebesgue în definirea și studierea integralei sale cu ajutorul măsurii Lebesgue. O dată adoptată această idee a lui Lebesgue, era inevitabil să se adopte și cealaltă idee a sa, relativă la înlocuirea diviziunilor orizontale prin diviziuni verticale. În felul acesta, s-a ajuns la o teorie de tip Lebesgue pentru integrala Riemann, teorie a cărei dezvoltare începe în jurul anului 1930. Această teorie trece, ca și teoria corespunzătoare a lui Lebesgue,

prin noțiunea de „funcție măsurabilă Jordan", noțiune care s-a cristalizat cu dificultate, după multe dibuiri. Am putea spune chiar că principalul efort, în construirea acestei teorii, a fost îndreptat către degajarea acelei clase de funcții care să îndeplinească aici rolul pe care funcțiile măsurabile Lebesgue îl au în teoria integralei Lebesgue. Un memoriu nefericit din 1929 al matematicianului olandez J. Ridder (de altfel, autor a numeroase contribuții în teoria integralei), care crezuse că noțiunea de funcție măsurabilă Jordan se poate obține din noțiunea de funcție măsurabilă Lebesgue prin simpla înlocuire a măsurabilității Lebesgue a mulțimilor prin măsurabilitatea lor Jordan, a dat un stimulent deosebit cercetărilor în această direcție, confirmîndu-se din nou cît de fecundă poate fi uneori o greșeală.

Cărei exigențe trebuia să-i răspundă definiția măsurabilității Jordan a unei funcții? Această definiție trebuia să fie de așa natură, încît orice funcție integrabilă Riemann pe $[a, b]$ să fie măsurabilă Jordan pe $[a, b]$ (întocmai după cum orice funcție integrabilă Lebesgue este măsurabilă Lebesgue) și orice funcție mărginită și măsurabilă Jordan pe $[a, b]$ să fie integrabilă Riemann pe $[a, b]$ (întocmai după cum orice funcție mărginită și măsurabilă Lebesgue pe $[a, b]$ este integrabilă Lebesgue pe $[a, b]$). Deoarece măsurabilitatea Lebesgue a unei funcții f revine la măsurabilitatea Lebesgue a mulțimilor $\{x; f(x) > \alpha\}$ pentru orice număr real α, Ridder definise (în memoriul amintit) măsurabilitatea Jordan a funcției f prin măsurabilitatea Jordan a aceleași mulțimi. Nu mică a fost mirarea cîțiva ani mai tîrziu (prin 1933) cînd, prin exemple date concomitent de către mai mulți matematicieni (unul dintre exemple fiind datorat marelui matematician român S. Stoilow), s-a constatat că, adoptîndu-se definiția propusă de Ridder, nici măcar funcțiile continue nu sînt, toate, măsurabile Jordan. Într-adevăr, fie E o mulțime de tip Cantor, dar de măsură Lebesgue pozitivă, construită pe $[0, 1]$. Dacă f este funcția egală, în fiecare punct x din $[0, 1]$, cu distanța de la x la mulțimea E, atunci f este continuă și nenegativă pe $[0, 1]$, fiind egală cu zero exact în punctele mulțimii E. Dar avem $l_i(E) = 0$ și $l_e(E) > 0$, deci mulțimea E nu este măsu-

rabilă Jordan. Rezultă că mulțimea $\{x: f(x) > 0\}$ (complementara lui E) nu este măsurabilă Jordan, deci funcția f, deși continuă, nu este măsurabilă Jordan în sensul propus de Ridder.

O versiune adecvată a noțiunii de funcție măsurabilă Jordan a fost obținută prin eforturile conjugate ale mai multor matematicieni (printre care și Ridder). Iată în ce constă această versiune: o funcție este, prin definiție, măsurabilă Jordan pe $[a, b]$, dacă mulțimea $\{x; f(x) > \alpha\}$ este măsurabilă Jordan pentru orice număr real α, cu excepția eventuală a unei mulțimi cel mult numărabile de valori ale lui α. Se demonstrează că o funcție reală f, de una sau mai multe variabile reale, este măsurabilă Jordan pe un interval I, dacă și numai dacă f este continuă aproape peste tot pe I. De aici rezultă imediat că o funcție mărginită și măsurabilă Jordan pe un interval compact este integrabilă Riemann pe acest interval și reciproc.

Cu ajutorul noțiunii de funcție măsurabilă Jordan putem construi o teorie corespunzătoare a integralei. Fiind dată o funcție f, definită, măsurabilă Jordan și mărginită pe o mulțime E măsurabilă Jordan, vom considera un interval $[A, B]$ care conține strict intervalul marginilor lui f $(A < m \leqslant M < B)$. Vom considera numai acele diviziuni Δ ale lui $[A, B]$ care sînt formate din puncte neaparținînd mulțimii numărabile excepționale din definiția măsurabilității Jordan a lui f. Fiecărei diviziuni $A = y_0 < \ldots < y_i < y_{i+1} < \ldots < y_n = B)$ îi vom asocia sumele

$$\sum_{i=0}^{n-1} y_i \, l(E_i), \quad \sum_{i=0}^{n-1} y_{i+1} l(E_i),$$

unde $E_i = \{x; y_i \leqslant f(x) < y_{i+1}\}$. De aici, mai departe, procedăm exact ca și în definiția integralei Lebesgue și obținem conceptul de integrabilitate Riemann și cel de integrală Riemann pe mulțimea măsurabilă Jordan E. Se demonstrează că orice funcție f, mărginită și măsurabilă Jordan pe E, este integrabilă Riemann. Tot odată, se arată că o funcție reală f este măsurabilă Jordan pe mulțimea E măsurabilă Jordan, dacă și numai dacă f este continuă aproape peste tot pe E (relativ la mulțimea E).

Aceste rezultate arată că singurul progres obținut (față de integrala Riemann clasică) prin adoptarea măsurii Jordan și a diviziunilor verticale este posibilitatea de a defini integrala Riemann pe mulțimi ceva mai generale decît intervalele, anume pe mulțimi măsurabile Jordan. În schimb, în ceea ce privește natura funcțiilor care se integrează, nu am progresat, rămînînd tot în cadrul funcțiilor mărginite și continue aproape peste tot.

Apare astfel clar că superioritatea integralei Lebesgue față de integrala Riemann nu provine — în esență — nici din adoptarea diviziunilor verticale, nici din adoptarea unei teorii a măsurii, ci din faptul că măsura adoptată de către Lebesgue este mai bună decît măsura Jordan. Prin ce este ea mai bună? O comparare atentă a celor două măsuri conduce imediat la concluzia că superioritatea măsurii Lebesgue față de măsura Jordan constă în faptul că în timp ce o reuniune numărabilă de mulțimi măsurabile Lebesgue este totdeauna măsurabilă Lebesgue, o reuniune numărabilă de mulțimi măsurabile Jordan nu este totdeauna o mulțime măsurabilă Jordan.

De altfel, teoria integralei Riemann pe mulțimi măsurabile Jordan poate fi construită și fără a se recurge la diviziuni verticale. Se pot folosi partiții finite ale mulțimii de definiție în mulțimi măsurabile Jordan și sumele corespunzătoare de tip Riemann.

Dacă însă mergem mai departe și definim, după metoda lui Lebesgue, integrabilitatea și integrala Riemann pentru funcții nemărginite, constatăm că nu mai regăsim integrala Riemann în sens generalizat. În timp ce aceasta din urmă admitea fenomenul de semiconvergență, noua integrală Riemann a unei funcții nemărginite nu mai admite acest fenomen. Este aici un nou avantaj pe care-l oferă o teorie de tip Lebesgue pentru integrala Riemann: dispare acea neconcordanța supărătoare dintre integrala Riemann în sens generalizat pentru funcții de o variabilă și aceeași integrală pentru funcții de mai multe variabile reale.

Trebuie sa mai observăm ca metoda lui Lebesgue a diviziunilor verticale aduce, chiar pentru integrala Riemann, un spor de eleganță și simplitate în desfășurarea

demonstrațiilor, ceea ce o recomandă și din punct de vedere didactic.

Să încercăm, în sfîrșit, să înțelegem de unde provin acele fenomene de continuitate pe care le detectăm atît în structura funcțiilor integrabile Riemann, cît și în aceea a funcțiilor integrabile Lebesgue. În definiția „orizontală" a integralei, pe fiecare mulțime A_i a partiției mulțimii de definiție în mulțimi măsurabile, funcția este reprezentată prin valoarea ei într-un punct ξ_i din A_i. Dar această reprezentare devine cu atît mai fidelă, cu cît sînt mai apropiate valorile lui f pe A_i de valoarea $f(\xi_i)$. Trebuie deci să ne așteptăm ca punctele care se sustrag acestui deziderat să nu poată ocupa un teritoriu prea mare. În definiția integralei Riemann mulțimile A_i sînt măsurabile Jordan, iar continuitatea pe A_i, prin A_i, revine aproape peste tot pe A_i la continuitatea obișnuită, deoarece o mulțime măsurabilă Jordan diferă de interiorul ei doar printr-o mulțime de măsură nulă. Așa se ajunge la continuitatea aproape peste tot a funcțiilor integrabile Riemann. În ceea ce privește integrala Lebesgue, aici mulțimile A_i sînt măsurabile Lebesgue. Se știe că aproape orice punct al unei mulțimi A măsurabile Lebesgue este un punct de densitate pentru A (a se vedea, pentru această noțiune, paragraful 5 din capitolul I). De aici rezultă că proprietatea de continuitate a unei funcții pe A, prin mulțimea A, revine la proprietatea de continuitate aproximativă aproape peste tot pe A (pentru proprietatea de continuitate aproximativă, a se vedea de asemenea paragraful 5 din capitolul I). Așa se ajunge la continuitatea aproximativă aproape peste tot a oricărei funcții integrabile Lebesgue.

INFIRMITĂȚI ALE INTEGRALEI LEBESGUE

Acum, după ce am adus atîtea laude integralei Lebesgue, să vedem totuși dacă ea rezolvă toate problemele în fața cărora integrala Riemann se dovedea insuficientă. Răspunsul, după cum se va vedea în acest paragraf, este negativ. Mai întîi, să ne amintim că suma unei serii trigo-

nometrice convergente pe $[a, b]$ nu este totdeauna integrabilă Riemann pe $[a, b]$. Unele dintre sumele de serii trigonometrice neintegrabile Riemann pe $[a, b]$ sînt integrabile Lebesgue pe $[a, b]$; din păcate însă, există serii trigonometrice convergente pe $[a, b]$, ale căror sume nu sînt integrabile Lebesgue pe $[a, b]$. De aici rezultă că nu orice serie trigonometrică convergentă este o serie Fourier-Lebesgue. Această situație a făcut necesară introducerea unei noi integrale, despre care vom discuta în paragraful următor.

O altă problemă importantă în care integrala Lebesgue și-a manifestat eficacitatea este aceea a exprimării primitivei unei funcții f (în cazul în care primitiva există) cu ajutorul integralei lui f. Integrala Riemann ne permitea să scriem relația

$$f(x) = f(a) + \int_a^x f'(t)\mathrm{d}t \tag{13}$$

ori de cîte ori derivata f' este integrabilă Riemann pe $[a, b]$. Integrala Lebesgue permite să se extindă valabilitatea relației (13) la toate funcțiile f cu derivată mărginită pe $[a, b]$. Mai mult decît atît, se poate arăta că relația (13) se menține pentru orice funcție f a cărei derivată există, finită, în fiecare punct din $[a, b]$ și este sumabilă pe $[a, b]$ (integrala fiind deci considerată în sensul lui Lebesgue). Astfel, relația (13) este adevărată, pe intervalul $[0, 1]$, pentru funcția

$$f(x) = \begin{cases} x^{3/2} \sin \dfrac{1}{x}, & \text{dacă } x > 0 \\ 0, & \text{dacă } x = 0 \end{cases}$$

a cărei derivată

$$f'(x) = \begin{cases} \dfrac{3}{2} x^{1/2} \sin \dfrac{1}{x} - x^{-1/2} \cos \dfrac{1}{x}, & \text{dacă } x > 0, \\ 0, & \text{dacă } x = 0 \end{cases}$$

este finită și sumabilă pe $[0, 1]$, deoarece are loc inegalitatea

$$|f'(x)| \leqslant \frac{3}{2} + \frac{1}{\sqrt{x}},$$

unde funcția din membrul al doilea este sumabilă pe [0, 1] dar $f'(x)$ nu este mărginită pe [0, 1]).

Totuși, există derivate finite pe [0, 1] care nu sînt sumabile Lebesgue pe [0, 1], Fie, de exemplu, funcția

$$f(x) = \begin{cases} x^2 \cos \frac{\pi}{x^2}, & \text{dacă } x > 0 \\ 0, & \text{dacă } x = 0. \end{cases}$$

Funcția aceasta este derivabilă în fiecare punct din [0, 1], dar derivata $f'(x)$, deși finită pe [0, 1], nu este sumabilă pe [0, 1]. Într-adevăr, dacă $0 < \alpha < \beta \leqslant 1$, atunci derivata f' este mărginită pe $[\alpha, \beta]$ și

$$\int_\alpha^\beta f'(x)\mathrm{d}x = \beta^2 \cos \frac{\pi}{\beta^2} - \alpha^2 \cos \frac{\pi}{\alpha^2}.$$

În particular, pentru

$$\alpha_n = \sqrt{\frac{2}{4n+1}}, \quad \beta_n = \frac{1}{\sqrt{2n}}$$

vom avea

$$\int_{\alpha_n}^{\beta_n} f'(x)\mathrm{d}x = \frac{1}{2n}.$$

Însă intervalele compacte $[\alpha_n, \beta_n]$ $(n = 1, 2, \ldots)$ sînt mutual disjuncte; aceasta înseamnă că punînd

$$E = \bigcup_{n=1}^{\infty} [\alpha_n, \beta_n],$$

vom avea

$$\int_E |f'(x)|\,\mathrm{d}x \geqslant \sum_{n=1}^{\infty} \frac{1}{2n} = +\infty,$$

și deci f' nu este sumabilă pe [0, 1]. Așadar, introducerea integralei Lebesgue nu rezolvă complet problema exprimării primitivei (în cazul în care ea există) unei funcții f cu ajutorul integralei lui f; cu alte cuvinte, folosind integrala Lebesgue, relația (13) nu va avea loc pentru orice funcție f cu derivată finită pe $[a, b]$, deoarece membrul al

doilea din (13) nu are sens pentru orice derivată finită $f'(x)$. Problema introducerii unui concept mai general de integrală, în raport cu care relația (13) să fie valabilă pentru orice funcție f cu derivata finită pe $[a, b]$, a fost rezolvată în deceniul al doilea al secolului nostru; soluția ei va fi expusă în unul din paragrafele următoare.

LA ESENȚA IDEII DE MĂSURĂ

Se poate pune întrebarea dacă nu cumva existența mulțimilor nemăsurabile Lebesgue nu influențează negativ teoria integralei Lebesgue, restrîngîndu-i cadrul. Într-adevăr, tocmai existența unor astfel de mulțimi face posibilă existența funcțiilor nemăsurabile Lebesgue și a mulțimilor pe care nu se poate defini integrala Lebesgue. Dacă nu ar exista mulțimi nemăsurabile Lebesgue, atunci orice funcție mărginită ar fi integrabilă Lebesgue pe orice mulțime mărginită.

Pentru a răspunde la această întrebare, trebuie să controlăm dacă existența mulțimilor nemăsurabile provine dintr-un defect, dintr-o particularitate a măsurii Lebesgue sau este ea inerentă ideii de măsură, așa cum ne este impusă de experiență. Ajungem astfel la necesitatea de a formula anumite deziderate pe care o funcție de mulțime trebuie să le satisfacă pentru ca ea să poată fi numită o măsură. Clasa $\mathcal{A}$ a mulțimilor pentru care vrem ca măsura să fie definită trebuie să constituie ceea ce se numește uneori un corp de mulțimi, adică să fie închisă față de operațiile de diferență și de reuniune numărabilă. În ceea ce privește măsura (pe care o notăm cu μ), ei i se cere să nu ia valori negative, să fie numărabil aditivă (adică măsura unei reuniuni numărabile de mulțimi disjuncte din $\mathcal{A}$ să fie egală cu suma seriei avînd ca termeni măsurile acestor mulțimi) și să fie finită pentru cel puțin o mulțime din $\mathcal{A}$. Dacă aceste condiții sînt îndeplinite, atunci se arată ușor că măsura mulțimii vide este egală cu zero (alegem o mulțime A de măsură finită; din $0 = A - A$ rezultă $\mu(0) = \mu(A) - \mu(A) = 0$) și că măsura este o funcție monotonă pe $\mathcal{A}$ (adică pentru două mulțimi A

și B din $\mathcal{A}$, A conținută în B, măsura lui A nu depășește măsura lui B; într-adevăr, $\mu(B) = \mu(A) + \mu(B - A)$, și, termenii fiind toți nenegativi, rezultă $\mu(A) \leqslant \mu(B)$).

Condițiile formulate mai sus constituie deziderate minimale pe care trebuie să le satisfacă o funcție de mulțime pentru ca ea să poată fi considerată o măsură, deci pentru ca ea să rețină ceea ce este esențial în ideea de măsură, așa cum se desprinde ea din nenumăratele situații particulare întîlnite în matematică, în celelalte științe și chiar în practica cotidiană. Pornindu-se de la această definiție sau de la definiții apropiate, s-a putut construi o teorie axiomatică a măsurii, ale cărei variante sînt expuse în unele cărți de specialitate (dintre care, în limba română: volumul III al tratatului de *Analiză matematică* al acad. Miron Nicolescu și manualul de *Teoria măsurii și funcții reale* al prof. N. Din[illegible]eanu). O astfel de teorie se va referi nu numai la măsuri pe corpuri de mulțimi ale dreptei numerice sau ale spațiului euclidian cu n dimensiuni, ci la corpuri de părți ale unei mulțimi abstracte. Un exemplu va fi edificator pentru generalitatea conceptului de măsură pe care tocmai l-am introdus: fie A o mulțime nevidă arbitrară. Familia tuturor părților lui A constituie, evident, un corp, fie el $\mathcal{A}$. Putem defini pe $\mathcal{A}$ o măsură în felul următor: alegem un element a din A și punem $\mu(E) = 1$, dacă E conține pe a și $\mu(E) = 0$ în cazul contrar. Se verifică ușor că toate proprietățile măsurii sînt satisfăcute. În particular, luînd în rolul mulțimii A chiar dreapta numerică, rezultă că se poate defini o măsură care să aibă sens pentru toate părțile dreptei numerice, deci în raport cu care să nu mai existe mulțimi nemăsurabile de puncte ale dreptei. S-ar părea deci că existența pe dreaptă a mulțimilor nemăsurabile Lebesgue se adaugă la lista infirmităților integralei Lebesgue.

Totuși, la o cercetare mai atentă, infirmitatea pe care tocmai am detectat-o se dovedește a fi doar aparentă. Într-adevăr, în definiția axiomatică a măsurii s-au neglijat unele deziderate intuitive care, în teoria integralei funcțiilor reale de o variabilă reală, s-au dovedit a fi foarte importante. Într-o teorie abstractă, axiomatică, se pornește de la un număr minim de restricții impuse obiectelor

cercetate și, treptat, cadrul se restrînge, se specializează prin formularea unor restricții suplimentare. În felul acesta se poate studia efectul fiecărei restricții cu care se lucrează, se poate stabili natura ei logică, poziția pe care ea o ocupă în ierarhia generală. Făcîndu-se, treptat, concesii în ceea ce privește generalitatea, putem în schimb să îmbogățim structura formațiunilor cercetate.

Aceste principii generale sînt folosite și în teoria măsurii. De la un corp abstract de mulțimi trecem la corpuri din ce în ce mai particulare, pentru a vedea ce specific capătă teoria măsurii în condițiile unor restricții mai numeroase. Printre aceste corpuri, un loc de seamă îl ocupă corpurile de părți ale spațiului euclidian n-dimensional. Aici, două restricții noi prezintă un deosebit interes. Prima constă în deziderатul — natural — ca printre mulțimile corpului considerat să se găsească toate intervalele, iar măsura acestor intervale să coincidă cu măsura lor în sens elementar (adică produsul dimensiunilor lor). De altfel, la această cerință am fost obligați să ne referim de mai multe ori în cursul acestui capitol. A doua restricție constă în cererea ca două mulțimi care se obțin una din alta printr-o deplasare destul de elementară să fie în același timp măsurabile sau nemăsurabile, iar în cazul în care sînt măsurabile, să aibă aceeași măsură. Ne-am întîlnit cu o proprietate de acest tip atunci cînd am demonstrat existența mulțimilor nemăsurabile Lebesgue. Am observat atunci faptul (care a intervenit efectiv în demonstrație) că măsurabilitatea și măsura Lebesgue, pentru mulțimi ale dreptei numerice, sînt invariante față de o translație.

Întorcîndu-ne acum la exemplul dat mai sus, de măsură definită pentru toate părțile dreptei numerice, observăm că această măsură nu este invariantă față de translații. Într-adevăr, punînd $\mu(E) = 1$, dacă E conține originea, și $\mu(E) = 0$, dacă E nu conține originea, măsura intervalului $[-1, +1]$ este egală cu 1, dar măsura intervalului $[1, 3]$, care se obține din primul printr-o translație de mărime 2, nu este egală cu 1, ci cu 0. Pe de altă parte, măsura unui interval care nu conține originea sau a unui interval care conține originea, dar nu este de lungime

egală cu 1 nu coincide cu lungimea acestui interval. Sînt astfel încălcate cele două condiții formulate mai sus. Problema naturală care se pune acum este aceea de a vedea dacă prin adoptarea acestor cerințe, ca proprietăți obligatorii ale măsurii, mai este posibil să definim o măsură pe corpul tuturor părților spațiului euclidian n-dimensional. Însă pentru ca problema să capete o formă precisă, va trebui în prealabil să precizăm clasa de deplasări în raport cu care măsurabilitatea și măsura vor trebui să fie invariante.

Fie φ o funcție definită pe spațiul euclidian n-dimensional R^n, cu valori în același spațiu. Vom spune că φ este o izometrie, dacă pentru orice pereche de puncte x și y ale spațiului distanța de la x la y coincide cu distanța de la $\varphi(x)$ la $\varphi(y)$. Ne vom referi, în cele ce urmează, numai la izometrii ale dreptei numerice în ea însăși.

Se constată ușor că o izometrie a dreptei este obligatoriu o aplicație biunivocă a dreptei pe ea însăși. În plus, se poate arăta că orice izometrie a dreptei numerice este de unul dintre următoarele două tipuri: $\varphi(x) = x + d$ sau $\varphi(x) = -x + d$ (unde d este un număr real). Se mai poate arăta că măsurabilitatea și măsura Lebesgue sînt invariante față de o izometrie.

Acum putem formula problema care ne preocupă într-o formă precisă:

Există o măsură definită pentru toate părțile dreptei numerice, invariantă față de orice izometrie a dreptei și reducîndu-se, în cazul intervalelor, la lungimea obișnuită?

Vom arăta că răspunsul la această întrebare este negativ (ceea ce constitue o reabilitare a măsurii Lebesgue în fața reproșului pe care i-l adresam, pentru mulțimile nemăsurabile la care ea conduce).

Să considerăm un șir $A_1, A_2, \ldots, A_n, \ldots$ de mulțimi mutual disjuncte, obținute una din alta prin cîte o izometrie și nemăsurabile Lebesgue, astfel încît

$$\left[-\frac{1}{2}, \frac{1}{2}\right] \subset \bigcup_{k=1}^{\infty} A_k \subset \left[-\frac{3}{2}, \frac{3}{2}\right].$$

(În obținerea mulțimilor A_k, cu proprietățile de mai sus, se folosește metoda aplicată în demonstrarea existenței mulțimilor nemăsurabile Lebesgue.)

Să presupunem acum, prin absurd, că există o măsură μ cu toate proprietățile enumerate mai sus. Vom avea

$$\mu\left(\left[-\frac{1}{2}, \frac{1}{2}\right]\right) \leqslant \sum_{k=1}^{\infty} \mu(A_k) \leqslant \mu\left(\left[-\frac{3}{2}, \frac{3}{2}\right]\right).$$

Dar intervalele $\left[-\frac{1}{2}, \frac{1}{2}\right]$ și $[0, 1]$ se obțin unul din celălalt printr-o translație; mulțimile A_k au aceeași măsură, fie ea λ, iar intervalul $\left[-\frac{3}{2}, \frac{3}{2}\right]$ este o mulțime mărginită. Rezultă că $1 \leqslant \lambda + \lambda + \ldots + \lambda + \ldots < +\infty$, ceea ce este imposibil atît pentru $\lambda > 0$, cît și pentru $\lambda = 0$. Am ajuns astfel la o contradicție care arată că existența unei măsuri cu proprietățile dorite nu este posibilă.

Așadar existența mulțimilor nemăsurabile Lebesgue are o rațiune profundă, decurgînd din natura complexă a cerințelor la care trebuia să răspundă măsura Lebesgue. În această ordine de idei, este semnificativă și următoarea teoremă a lui Banach și Kuratowski: Admițînd ipoteza continuului (adică inexistența unei mulțimi nenumărabile de numere reale, care nu este echivalentă cu mulțimea tuturor numerelor reale), nu există nici o măsură neidentic nulă definită pentru toate părțile intervalului $[0, 1]$ și anulîndu-se pentru orice mulțime care se reduce la un punct. Să spunem, cu acest prilej, că pe baza rezultatelor lui Gödel și Cohen ipoteza continuului este, ca și axioma alegerii, o propoziție independentă.

INTEGRALELE LUI DENJOY ȘI INTEGRALA LUI PERRON

În 1912, matematicianul francez Arnaud Denjoy a introdus o integrală care avea să se dovedească superioară integralei Lebesgue cel puțin din punctele de vedere discutate în penultimul paragraf: integrarea sumelor de serii trigonometrice și integrarea funcțiilor derivate. Integrala lui Denjoy se definește printr-un proces complicat de inducție transfinită, pe care nu-l putem expune aici. Vom prezenta însă o altă integrală, introdusă de către matematicianul

german Oscar Perron în 1914, integrală despre care s-a constatat ulterior (G. Hacke 1921, P. S. Alexandrov 1924, G. Looman 1925) că este echivalentă cu integrala Denjoy: O funcție reală definită pe $[a, b]$ este integrabilă Denjoy pe $[a, b]$, dacă și numai dacă ea este integrabilă Perron pe $[a, b]$, iar valorile celor două integrale sînt egale. În același timp, orice funcție sumabilă pe $[a, b]$ este integrabilă Denjoy și Perron pe $[a, b]$ iar cele trei integrale sînt egale.

Fie F o funcție reală definită pe $[a, b]$ și fie x_0 un punct din $[a, b]$. Numerele

$$\underline{D}F(x_0) = \liminf_{x \to x_0} \frac{F(x) - F(x_0)}{x - x_0}, \quad \overline{D}F(x_0) = \limsup_{x \to x_0} \frac{F(x) - F(x_0)}{x - x_0}$$

se numesc — respectiv — derivata inferioară și derivata superioară a funcției F în punctul x_0.

Fie f o funcție reală definită pe intervalul $[a, b]$ și putînd lua, eventual, și valori infinite. O funcție reală F, continuă pe $[a, b]$, este o semiprimitivă superioară a funcției f pe $[a, b]$, dacă următoarele trei condiții sînt îndeplinite:

1) $F(a) = 0$; 2) $\underline{D}F(x) > -\infty$ pentru orice x din $[a, b]$;

3) $\underline{D}F(x) \geqslant f(x)$ pentru orice x din $[a, b]$.

Funcția reală F, continuă pe $[a, b]$, este o semiprimitivă inferioară a funcției f pe $[a, b]$, dacă sînt îndeplinite următoarele condiții: 1) $F(a) = 0$; 2) $\overline{D}F(x) < +\infty$ pentru orice x din $[a, b]$; 3) $\overline{D}F(x) \leqslant f(x)$ pentru orice x din $[a, b]$.

Noțiunile de semiprimitivă inferioară și semiprimitivă superioară sînt extensiuni ale noțiunii de primitivă: dacă funcția reală f este derivata funcției F pe intervalul $[a, b]$ și dacă $F(a) = 0$, atunci F este, în același timp, o semiprimitivă inferioară și o semiprimitivă superioară a funcției f pe $[a, b]$.

Vom spune că funcția reală f, definită pe $[a, b]$, este integrabilă Perron pe $[a, b]$, dacă următoarele două condiții sînt îndeplinite:

α) funcția f admite pe $[a, b]$ cel puțin o semiprimitivă inferioară și cel puțin o semiprimitivă superioară;

β) marginea inferioară a valorilor luate de semiprimitivele superioare ale lui f în punctul b este egală cu

marginea superioară a valorilor luate de semiprimitivele inferioare ale lui f în același punct b.

Dacă funcția f este integrabilă, numim valoarea comună a celor două margini despre care este vorba în condiția β integrala Perron a funcției f pe $[a, b]$.

Vom da acum trei leme, pe baza cărora vom putea arăta că orice derivată finită este integrabilă Perron și are loc, pentru ea, o formulă de tip Leibniz-Newton.

Lema 1. Fie U și V două funcții finite pe $[a, b]$. Dacă pentru un punct x_0 din $[a, b]$ avem $\underline{D}U(x_0) > -\infty$, $\overline{D}V(x_0) < +\infty$, atunci, notînd $R(x) = U(x) - V(x)$, avem $\underline{D}R(x_0) \geqslant \underline{D}U(x_0) - \overline{D}V(x_0)$.

Demonstrație. Fie $\{h_k\}$ un șir pentru care $h_k \neq 0$ $(k = 1, 2, \ldots)$, $h_k \to 0$ și

$$\lim \frac{R(x_0 + h_k) - R(x_0)}{h_k} = DR(x_0).$$

Rezultă existența limitelor

$$\lambda = \lim \frac{U(x_0 + h_k) - U(x_0)}{h_k}, \qquad \mu = \lim \frac{V(x_0 - h_k) - V(x_0)}{h_k}.$$

În baza inegalităților din ipoteză, avem $\lambda > -\infty$ și $\mu < +\infty$, deci are sens diferența $\lambda - \mu$, și avem $\underline{D}R(x_0) = \lambda - \mu$. Lema 1 rezultă acum imediat, observînd că $\lambda \geqslant \underline{D}U(x_0)$ și $\mu \leqslant \overline{D}V(x_0)$.

Lema 2. Dacă U este o semiprimitivă superioară, iar V este o semiprimitivă inferioară a funcției f pe $[a, b]$, atunci diferența $R(x) = U(x) - V(x)$ este o funcție monoton crescătoare.

Demonstrație. În baza lemei 1, avem pentru orice x din $[a, b]$: $\underline{D}R(x) \geqslant \underline{D}U(x) - \overline{D}V(x) \geqslant 0$. Însă se poate arăta că dacă toate numerele derivate ale unei funcții reale, definite pe $[a, b]$), sînt nenegative pe $[a, b]$, atunci funcția în cauză este monoton crescătoare. Observăm că funcția R satisface aceste premise, deci ea este monoton crescătoare pe $[a, b]$.

Corolarul 1. În condițiile lemei 2, avem $U(b) \geqslant V(b)$.

Lema 3. Dacă funcția reală f este derivata funcției F pe $[a, b,]$ și dacă $F(a) = 0$, atunci numărul $F(b)$ este cel mai mic dintre numerele $U(b)$ și cel mai mare dintre

numerele $V(b)$ (unde U parcurge toate semiprimitivele superioare, iar V toate semiprimitivele inferioare ale lui f pe $[a, b]$).

Demonstrație. Lema 3 este o consecință imediată a faptului că F este, pentru f, atît semiprimitivă inferioară, cît și semiprimitivă superioară.

Putem trece acum la enunțarea rezultatului anunțat, observînd că el este o consecință a lemelor 2 și 3 de mai sus:

Dacă funcția F admite în fiecare punct din $[a, b]$ o derivată finită f, atunci f este integrabilă Perron pe $[a, b]$, iar integrala Perron a funcției f pe $[a, b]$ este egală cu diferența valorilor lui F în punctele b și a.

Atragem atenția asupra faptului că în teorema de mai sus nu se poate renunța la ipoteza de finitudine a funcției f (după cum a arătat V. I. Kozlov în 1951). Există derivate care devin infinite chiar într-o infinitate de puncte, dar se poate arăta că această infinitate nu poate fi decît de măsură nulă. Pătrundem astfel într-o zonă în care nici integrala Perron nu mai poate face față scopului propus, acela al validării formulei lui Leibniz și Newton.

După cum am afirmat mai sus, integrala introdusă de Denjoy în 1912 este echivalentă cu integrala Perron. Din acest motiv, ea a primit numele de integrala Denjoy-Perron. Este interesantă reacția lui Denjoy la demonstrarea echivalenței dintre integrala sa și integrala lui Perron. Denjoy s-a arătat nemulțumit de metoda lui Perron, obiectîndu-i caracterul ei neefectiv. El observa că dacă se cere un avion cu o anumită viteză k, sînt posibile două răspunsuri: construirea avionului respectiv (aceasta ar fi metoda lui Denjoy) sau, ordonînd avioanele după viteza lor, considerarea marginii inferioare a avioanelor cu viteză superioară lui k (aceasta ar fi metoda lui Perron!). În fapt, cele două metode s-au dovedit a fi egal de utile în analiză, stimulînd fiecare dezvoltarea ulterioară a teoriei integralei.

Să precizăm acum natura raporturilor dintre integrala Lebesgue și integrala Denjoy-Perron, încercînd apoi să observăm cît de mult se poate deosebi structura unei funcții integrabile Perron de aceea a unei funcții sumabile.

Dacă funcția f este sumabilă pe $[a, b]$, atunci ea este integrabilă Perron pe $[a, b]$ și cele două integrale sînt egale.

Din acest rezultat se poate deduce imediat un altul, deosebit de semnificativ. Fie $U(x)$ o semiprimitivă superioară a funcției f, presupusă integrabilă Perron și nenegativă pe $[a, b]$. Din inegalitățile $\underline{D}U(x) \geqslant f(x) \geqslant 0$ rezultă că $U(x)$ este o funcție monoton crescătoare și deci derivata ei (care există în orice caz aproape peste tot) este sumabilă pe $[a, b]$. Ținînd seama că în fiecare punct în care U admite derivată, avem $0 \leqslant f(x) \leqslant U'(x)$, rezultă:

Orice funcție nenegativă și integrabilă Perron pe $[a, b]$ este sumabilă pe $[a, b]$.

Cu alte cuvinte, superioritatea integralei Perron în raport cu integrala Lebesgue se manifestă numai în ceea ce privește funcțiile de semn variabil. În particular, rezultă că orice derivată finită și nenegativă este sumabilă. O altă consecință importantă a rezultatului obținut este aceea că fenomenul de convergență absolută nu poate apărea la integrala Perron decît atunci cînd aceasta există și ca integrală Lebesgue. Integrala Denjoy-Perron, în ceea ce are ea specific, este semiconvergentă.

Un rezultat oarecum dezamăgitor, în sensul că dezvăluie caracterul destul de limitat al progresului realizat de integrala Perron, este cel care afirmă măsurabilitatea Lebesgue a oricărei funcții integrabile Perron pe $[a, b]$. Într-adevăr, notînd cu $F(x)$ integrala Perron a funcției f de la a la x și prelungind funcția F la dreapta lui b, punînd $F(x) = F(b)$ pentru $x > b$, observăm că aproape pentru orice x din $[a, b]$ avem

$$f(x) = \lim_{n \to \infty} n \left[F\left(x + \frac{1}{n}\right) - F(x) \right].$$

Însă se poate arăta că F este continuă pe $[a, b]$, deci f este, aproape peste tot pe $[a, b]$, limita unui șir de funcții continue. Așadar f este măsurabilă Lebesgue pe $[a, b]$.

După cum am văzut, cu integrala Denjoy-Perron se poate integra orice derivată finită. Însă în prezentarea conceptului de funcție măsurabilă Lebesgue se constată că tipul de derivabilitate legat în mod natural de măsura-

bilitatea Lebesgue nu este derivabilitatea obișnuită, ci derivabilitatea aproximativă (sau asimptotică). Derivabilitatea aproximativă a funcției F în punctul x_0 revine la existența și finitudinea limitei raportului

$$\frac{F(x) - F(x_0)}{x - x_0}$$

printr-o mulțime care admite punctul x_0 ca punct de densitate. (Noțiunea de punct de densitate al unei mulțimi a fost prezentată în paragraful (5) din capitolul I). Constatîndu-se că nu orice derivată aproximativă este integrabilă Denjoy-Perron, s-a simțit nevoia unei integrabilități mai generale. Ca urmare a acestui fapt, în 1916 Arnaud Denjoy și, independent de el, A. I. Hincin introduc o integrală mai generală decît integrala Denjoy-Perron. Noua integrală a primit denumirea de integrala Denjoy-Hincin. Cu ajutorul ei se poate integra orice derivată aproximativă finită.

Dacă definiția constructivă a integralelor Denjoy-Perron și Denjoy-Hincin este prea complicată din punct de vedere tehnic pentru a o putea prezenta aici, ne va fi în schimb mai ușor să prezentăm definiția lor descriptivă, echivalentă cu cea constructivă. Această definiție are și avantajul de a da o idee mai precisă asupra deosebirii dintre funcțiile care sînt integrale Lebesgue nedefinite și funcțiile care sînt integrale Denjoy nedefinite. Vom vedea că există extensiuni ale proprietății de continuitate absolută care caracterizează integralele Denjoy nedefinite exact așa cum continuitatea absolută caracterizează integralele Lebesgue nedefinite. O funcție reală f, de o variabilă reală, este, prin definiție, absolut continuă pe o mulțime E, dacă pentru orice $\varepsilon > 0$ există un $\eta > 0$, cu proprietatea că oricare ar fi șirul de intervale $\{[a_k, b_k]\}$ fără puncte interioare comune două cîte două, dar ale căror extremități a_k, b_k aparțin lui E, inegalitatea

$$\sum_k (b_k - a_k) < \eta$$

implică inegalitatea

$$\sum_k |f(b_k) - f(a_k)| < \varepsilon.$$

Dacă o funcție reală F, de o variabilă reală, este continuă pe mulțimea E, iar E este o reuniune cel mult numărabilă de mulțimi E_n ($n = 1, 2, \ldots$), cu proprietatea că F este absolut continuă pe fiecare mulțime E_n ($n = 1, 2, \ldots$), atunci spunem că F este absolut continuă în sens generalizat pe mulțimea E.

Se poate arăta că orice funcție absolut continuă în sens generalizat pe o mulțime E măsurabilă Lebesgue este aproximativ derivabilă aproape peste tot pe E. Orice funcție absolut continuă pe un interval $[a, b]$ este absolut continuă în sens generalizat pe $[a, b]$, dar reciproca nu este adevărată.

Acum putem introduce pe o cale descriptivă integrabilitatea și integrala Denjoy-Hincin.

Spunem că o funcție reală f definită aproape peste tot pe $[a, b]$ este integrabilă Denjoy-Hincin pe $[a, b]$, dacă există o funcție F, absolut continuă în sens generalizat pe $[a, b]$, astfel încît derivata aproximativă a lui F să fie egală — aproape peste tot pe $[a, b]$ — cu f. Diferența $F(b) - F(a)$ se numește integrala Denjoy-Hincin a funcției f pe $[a, b]$. Valoarea acestei integrale nu depinde de alegerea funcției F cu proprietatea indicată, deoarece se poate arăta că dacă două funcții absolut continue în sens generalizat au aproape peste tot pe $[a, b]$ aceeași derivată aproximativă, atunci cele două funcții diferă printr-o constantă.

Pentru a introduce pe cale descriptivă integrabilitatea și integrala Denjoy-Perron, vom defini, în prealabil, o altă extensiune a proprietății de continuitate absolută, extensiune mai puțin largă decît cea folosită în definiția integralei Denjoy-Hincin.

O funcție reală g de o variabilă reală este, prin definiție, absolut continua în sens restrîns pe mulțimea mărginită E, dacă g este mărginită pe un anumit interval conținînd pe E și daca, în plus, există pentru orice $\varepsilon > 0$ un $\eta > 0$, astfel încît, oricare ar fi șirul $\{I_k\}$ de intervale fără puncte interioare comune două cîte două și ale căror extremități aparțin lui E, inegalitatea

$$\sum_k |I_k| < \eta$$

(unde prin $|I_k|$ s-a notat lungimea lui I_k), implică inegalitatea

$$\sum_k \omega(g;\ I_k) < \varepsilon,$$

unde $\omega(g;\ I_k)$ reprezintă oscilația funcției g pe intervalul I_k.

O funcție reală G de o variabilă reală este, prin definiție, absolut continuă în sens cvasigeneralizat pe mulțimea E, dacă G este continuă pe E și dacă E se poate exprima ca o reuniune cel mult numărabilă de mulțimi mărginite E_n $(n = 1, 2, \ldots)$, astfel încît G este absolut continuă în sens restrîns pe fiecare E_n $(n = 1, 2, \ldots)$.

Se poate arăta că orice funcție absolut continuă în sens cvasigeneralizat pe $[a, b]$ este derivabilă aproape peste tot pe $[a, b]$. O funcție absolut continuă pe $[a, b]$ este absolut continuă în sens cvasigeneralizat pe $[a, b]$, dar reciproca nu este adevărată.

Fiind dată o funcție reală f definită aproape peste tot pe $[a, b]$, spunem că f este integrabilă Denjoy-Perron pe $[a, b]$, dacă există o funcție reală F, absolut continuă în sens cvasigeneralizat pe $[a, b]$, astfel încît, aproape peste tot pe $[a, b]$, derivata lui F este egală cu f. Diferența $F(b)—F(a)$ se numește integrala Denjoy-Perron a funcției f pe $[a, b]$; ea este unic determinată, deoarece două funcții absolut continue în sens cvasigeneralizat, care au aproape peste tot pe $[a, b]$ aceeași derivată, diferă printr-o constantă.

DE LA STIELTJES LA RIESZ. O NOUĂ PERSPECTIVĂ ÎN TEORIA INTEGRALEI

La sfîrșitul secolului trecut, matematicianul francez, de origine olandeză, Th. Stieltjes a imaginat o extensiune a integralei Riemann, care cîtva timp a trecut neobservată, pînă ce matematicianul maghiar Fr. Riesz a adus-o în centrul atenției, la sfîrșitul primului deceniu al secolului nostru, printr-un rezultat surprinzător.

Unele probleme din teoria fracțiilor continue, precum și unele probleme de mecanică, cum ar fi aceea a definirii

și determinării momentului static al unei bare, l-au condus pe Stieltjes la un proces de integrare a unei funcții f în raport cu o altă funcție g, ambele funcții fiind definite pe un același interval $[a, b]$. Integrabilitatea și integrala Stieltjes se definesc ca și integrabilitatea și integrala Riemann, cu singura deosebire că în locul sumelor (5) (a se vedea paragraful 12 al acestui capitol) se consideră sumele

$$\sigma_\Delta(f, g, \xi_i) = \sum_{i=0}^{n-1} f(\xi_i)(g(x_{i+1}) - g(x_i)), \qquad (14)$$

unde $x_i \leqslant \xi_i \leqslant x_{i+1}$ pentru $0 \leqslant i \leqslant n-1$. După cum se vede, sumele (5) sînt acel caz particular al sumelor (14) în care funcția g este chiar aplicația identică a intervalului $[a, b]$. Sumele (14) manifestă însă o sensibilitate incomparabil mai mare decît sumele (5). Să considerăm, de pildă, următoarele trei restricții: α) $\xi_i = x_i$ pentru $0 \leqslant i \leqslant n-1$ (această restricție, impusă sumelor (5), conduce la sumele considerate de (Cauchy); β) $\xi_i = x_{i+1}$ pentru $0 \leqslant i \leqslant n-1$; γ) $x_i < \xi_i < x_{i+1}$ pentru $0 \leqslant i \leqslant n-1$. Nici una dintre restricțiile α, β și γ, aplicate sumelor (5), nu se repercutează asupra tipului de integrabilitate și de integrală obținut; se poate arăta că de fiecare dată se obține integrabilitatea și integrala Riemann. În schimb, fiecare dintre restricțiile α, β și γ, aplicate sumelor (14), conduce la un alt tip, la o altă noțiune de integrabilitate Stieltjes.

Să introducem acum un tip nou de trecere la limită pentru sumele integrale, considerate ca funcții de diviziunea Δ. Vom spune că sumele σ_Δ au ca „limită după finețe" numărul I, dacă fiecărui număr pozitiv ε îi corespunde o diviziune Δ' a lui $[a, b]$, astfel încît pentru orice diviziune Δ mai fină decît Δ' să avem $|\sigma_\Delta - I| < \varepsilon$. Se poate arăta că luînd limita după finețe a sumelor (5), obținem tot integrabilitatea și integrala Riemann; acest lucru rămîne adevărat chiar dacă impunem sumelor (5) una oarecare dintre restricțiile α, β și γ, definite mai sus. În schimb, limita după finețe a sumelor (14), luată fără nici o restricție asupra valorilor sau combinată cu una dintre restricțiile α, β, γ, conduce la patru tipuri noi de integrabilitate Stieltjes. Am inventariat deci, pînă în

momentul de față, opt tipuri distincte de integrabilitate Stieltjes, dar aceste opt tipuri se reduc, în cazul în care g este aplicația identică, la unul și același tip de integrabilitate: cea în sensul lui Riemann.

Acest contrast între „insensibilitatea" sumelor (5) și susceptibilitatea deosebită a sumelor (14) nu se oprește aici. Se știe că la integrala Riemann se poate ajunge și prin considerarea sumelor Darboux, adică a acelor sume care se obțin din (5), înlocuind pe $f(\xi_i)$ prin m_i, respectiv M_i (unde m_i și M_i sînt marginile inferioară și superioară ale funcției f pe intervalul $[x_i, x_{i+1}]$). În schimb, presupunînd că funcția f este mărginită pe $[a, b]$ și că funcția g este crescătoare pe $[a, b]$ și înlocuind sumele (14) prin sumele Darboux-Stieltjes corespunzătoare, se obține un tip de integrabilitate Stieltjes diferit de cel obișnuit. Combinînd considerarea sumelor Darboux-Stieltjes cu restricțiile α, β, γ și adoptînd, alternativ, pentru sumele σ_Δ limita obișnuită (numită și limita după normă) și limita după finețe, obținem un mare număr de tipuri noi de integrabilitate Stieltjes. Trebuie să spunem că multe dintre tipurile de integrabilitate Stieltjes obținute pe căile indicate mai sus intervin efectiv în diferite capitole ale matematicii. Astfel, pentru a da un singur exemplu, integrabilitatea Stieltjes definită cu ajutorul sumelor (14), neimpunînd nici una dintre restricțiile α, β, γ, dar adoptînd limita după finețe, este fundamentală în teoria probabilităților.

Rezultatul surprinzător obținut de Riesz se referă la tipul de integrabilitate Stieltjes definit cu ajutorul sumelor (14), fără a se impune vreo restricție suplimentară valorilor ξ_i și adoptîndu-se limita după normă. O teoremă elementară afirmă că orice funcție f continuă pe $[a, b]$ este integrabilă Stieltjes în raport cu orice funcție g cu variația mărginită pe $[a, b]$. Fiind dată o funcție g cu variație mărginită pe $[a, b]$, să notăm cu $F(f)$ integrala Stieltjes a funcției f, continue pe $[a, b]$, în raport cu g. F este deci o aplicație definită pe mulțimea funcțiilor continue pe $[a, b]$ și cu valori reale. Se arată ușor că există o constantă k cu proprietatea

$$|F(f)| \leqslant k \cdot M(f), \tag{15}$$

unde prin $M(f)$ s-a notat maximul lui f pe $[a, b]$. (Este suficient să luăm pe k egal cu variația totală a lui g de la a la b.) Pe de altă parte, fiind date două funcții f_1 și f_2, continue pe $[a, b]$, avem, cum se verifică ușor.

$$F(f_1 + f_2) = F(f_1) + F(f_2). \tag{16}$$

Orice aplicație F a mulțimii funcțiilor continue pe $[a, b]$ în dreapta numerică, avînd proprietățile (15) și (16), constituie, prin definiție, o funcțională liniară pe spațiul $\mathcal{C}$ al funcțiilor continue. Rezultă deci că fiecare funcție g, cu variație mărginită pe $[a, b]$, induce o astfel de funcțională liniară F, funcțională care asociază fiecărei funcții f, continue pe $[a, b]$, integrala Stieltjes a acestei funcții în raport cu g. Rezultatul remarcabil al lui Riesz constă în descoperirea faptului că funcționalele liniare pe $\mathcal{C}$, induse de funcții cu variație mărginită, epuizează toate funcționalele liniare existente pe $\mathcal{C}$. Cu alte cuvinte, fiind dată o aplicație F a lui $\mathcal{C}$ în R, cu proprietățile (15) și (16), există o funcție g cu variație mărginită pe $[a, b]$, astfel încît pentru orice funcție f continuă pe $[a, b]$, $F(f)$ este tocmai integrala Stieltjes a lui f în raport cu g pe $[a, b]$.

Se deschide astfel o perspectivă cu totul nouă, atît în teoria integralei Stieltjes, cît și în aceea a funcționalelor liniare. Pe de o parte, orice funcțională liniară pe $\mathcal{C}$ admite o reprezentare integrală, pe de altă parte, integrala Stieltjes poate fi definită ca o funcțională liniară. Acest punct de vedere a contaminat întreaga dezvoltare ulterioară a teoriei integralei. În particular, chiar pentru diferitele tipuri de integrale Stieltjes, despre care s-a vorbit mai sus, s-au studiat funcționalele liniare asociate. Aici, apar pe primul plan dependența integralei de funcția care se integrează, proprietățile algebrice ale acestei dependențe și modul de organizare algebrică a mulțimii funcțiilor integrabile. Dar cu aceasta pătrundem într-o etapă nouă a dezvoltării analizei, etapă în care studiul funcțiilor individuale este înlocuit prin studiul claselor de funcții și al diferitelor aplicații definite pe aceste clase.

V

Ce se întîmplă dincolo de mulțimile și funcțiile boreliene?

INTRODUCERE

La 8 ianuarie 1967 s-au împlinit cincizeci de ani de de cînd matematicianul francez Jacques Hadamard a prezentat la Academia de Științe din Paris notele, care aveau să devină celebre, a doi matematicieni din Moscova, Mihail Suslin și Nikolae Luzin. În aceste note se semnala existența unei clase de mulțimi necunoscute pînă atunci, numite de descoperitorii lor „mulțimi analitice". Teoria acestor mulțimi n-a încetat, pînă astăzi, să pună matematicienilor probleme pasionante, dar deosebit de grele.

În capitolul de față vom prezenta unele aspecte ale acestei teorii destul de puțin cunoscută la noi în țară, mai cu seamă de către matematicienii tineri. Unele rezultate remarcabile referitoare la mulțimile analitice au fost publicate în revista „Mathematica" de la Cluj.

Noțiunea de mulțime analitică a fost introdusă de Mihail Suslin în 1917, iar noțiunea de mulțime proiectivă a fost introdusă de Nikolae Luzin în 1924. Teoria acestor mulțimi este de o mare profunzime. În special, mulțimile proiective prezintă proprietăți atît de surprinzătoare, încît Luzin a pus însăși problema legitimității lor.

Multe din problemele legate de aceste mulțimi sînt atît de grele, încît ele se află astăzi, după aproape cincizeci de ani de la apariție, aproape în același stadiu ca atunci cînd au fost puse. Matematicienii nu le-au putut încă rezolva. Paralel cu aceasta, dezvoltarea matematicii în ultimele decenii a scos de nenumărate ori în evidență importanța mulțimilor analitice și proiective, modul

natural în care ele apar în cele mai variate probleme de teoria funcţiilor de variabilă reală, topologie, analiză funcţională şi chiar în teoria funcţiilor de variabilă compexă.

O GREŞEALĂ BINE INSPIRATĂ

Descoperirea de către Suslin a mulţimilor analitice este legată de o întîmplare puţin obişnuită pentru o mare descoperire. Suslin era elevul lui Luzin. Luzin, în afară de faptul că era un mare om de ştiinţă, era şi mare animator. El a avut mulţi elevi şi a determinat, în bună măsură, direcţia de dezvoltare a şcolii matematice din Moscova. De altfel, se ştie că şcoala sovietică de topologie, ai cărei pionieri sînt P. S. Aleksandrov şi P. Urîson, este născută din preocupările de teoria funcţiilor reale iniţiate de Egorov şi mai ales de Luzin.

Luzin a remarcat repede talentul matematic al lui Suslin şi i-a recomandat să studieze celebrul memoriu al lui Lebesgue: *Sur les fonctions représentables analytiquement*, memoriu care pune multe probleme („Journal de mathématiques pures et appliquées", 1905).

Într-o zi, Suslin se prezintă la profesorul său, spunîndu-i că a găsit o greşeală în memoriul lui Lebesgue. Matematicianul polonez Waclaw Sierpiński, care a fost martor la această scenă, povesteşte că Luzin l-a ascultat cu multă atenţie pe elevul său, în ciuda faptului că era puţin plauzibil ca el într-adevăr să fi găsit o greşeală în memoriul unui matematician atît de mare ca Lebesgue. Observaţia lui Suslin s-a dovedit justă. Pentru a înţelege în ce constă ea, trebuie să spunem mai întîi că funcţiile reprezentabile sau exprimabile analitic, pe care le consideră Lebesgue în memoriul său, coincid cu funcţiile boreliene sau, ceea ce este tot una, cu funcţiile din clasificaţia lui Baire (aceste noţiuni vor fi definite mai jos). Lebesgue enunţă şi încearcă să demonstreze următoarea teoremă: „O funcţie definită implicit cu ajutorul unor reprezentări analitice este exprimabilă analitic în mod explicit". Particularizînd, dar făcîndu-l mai clar, acest enunţ revine

la următorul: „Dacă $x = f(t)$ este o funcție reprezentabilă analitic și inversabilă, atunci funcția inversă $t = \varphi(x)$ este de asemenea reprezentabilă analitic". Demonstrația lui Lebesgue decurge astfel: „Datorită ipotezei, se poate arăta că mulțimea punctelor din planul coordonatelor t, x în care $x - f(t) = 0$ este o mulțime boreliană E. Deci oricare ar fi $\alpha < \beta$, partea B a lui E pentru care $\alpha \leqslant t \leqslant \beta$ este o mulțime boreliană. Însă proiecția, pe axa absciselor, a unui interval bidimensional este un interval unidimensional, proiecția unei reuniuni de mulțimi este reuniunea proiecțiilor acestor mulțimi, iar proiecția intersecției unui șir descrescător de mulțimi este egală cu intersecția proiecțiilor mulțimilor din șir. Ținînd seamă că mulțimile boreliene coincid cu mulțimile care se obțin pornind de la intervale, prin aplicarea repetată a operațiilor de reuniune numărabilă și intersecție de șiruri descrescătoare, rezultă că proiecția mulțimii B pe axa $t = 0$ este o mulțime boreliană. Prin aceasta s-a demonstrat că funcția $t = \varphi(x)$ este boreliană".

Nu este greu de găsit greșeala în această demonstrație. Afirmația că proiecția intersecției unui șir descrescător de mulțimi este egală cu intersecția proiecțiilor acestor mulțimii este falsă.

Într-adevăr, fie E_n mulțimea tuturor punctelor din plan, ale căror coordonate x, y satisfac condițiile $x = 0$, $0 < y < \frac{1}{n}$. Este vizibil că nu există nici un punct comun tuturor mulțimilor $E_n (n = 1, 2, \ldots)$, deci proiecția pe axa absciselor a mulțimii $\bigcap_{n=1}^{\infty} E_n$ este mulțimea vidă. Pe de altă parte, oricare ar fi n, proiecția pe axa absciselor a mulțimii E_n este nevidă, fiind formată dintr-un punct: originea. Deci intersecția proiecțiilor mulțimilor E_n $(n = 1, 2, \ldots,)$ pe axa absciselor este nevidă.

Important nu este faptul că Suslin a găsit o greșeală în memoriul lui Lebesgue. Pentru Lebesgue, această greșeală era explicabilă din punct de vedere psihologie, ea fiind legată de o chestiune laterală față de scopul prin-

cipal al memoriului său. Greșeala este cu totul locală și nu atinge cu nimic valoarea acestui memoriu. Intuiția lui Lebesgue a fost, de altfel, justă. Într-adevăr, ulterior Luzin a dat o demonstrație corectă teoremei lui Lebesgue, Această demonstrație se sprijină în mod esențial pe teoria mulțimilor analitice și, pînă astăzi, nu se cunoaște o posibilitate de a demonstra teorema lui Lebesgue fără a se recurge la mulțimi analitice.

Ceea ce este important este faptul că Suslin, analizînd pînă la capăt greșeala lui Lebesgue, a descoperit că o operație atît de simplă, cum este proiecția, aplicată unei mulțimi boreliene, poate să conducă la o mulțime neboreliană. Mulțimile analitice liniare nu sînt altceva decît mulțimile liniare care se obțin ca proiecții ale mulțimilor boreliene din plan.

Din păcate, Suslin a murit foarte tînăr (în 1918). Luzin este cel care a construit întreaga teorie a mulțimilor analitice, cu prelungirea ei firească, teoria mulțimilor proiective. Această teorie a lui Luzin, împreună cu contribuția inițială a lui Suslin, se află expusă într-o carte celebră (a lui Luzin, publicată la Paris, în limba franceză): *Lecții despre mulțimile analitice și aplicațiile lor.* În 1953 a apărut la Moscova o ediție nouă a acestei cărți, ediție îngrijită de doi dintre cei mai valoroși continuatori ai lui Luzin, Ludmila Keldîș și P. S. Novikov (Luzin a murit în 1950). Această ediție conține, sub formă de completări și observații, unele contribuții care s-au adus în teoria mulțimilor analitice și proiective după apariția primei ediții a acestei cărți, în 1930.

Luzin, în modestia sa, merge pînă acolo încît atribuie lui Lebesgue meritul de a fi descoperit mulțimile analitice, deoarece (după cum vom arăta) construcția acestor mulțimi este implicit conținută în memoriul lui Lebesgue. Dar Lebesgue, în prefața pe care a scris-o la ediția franceză a cărții lui Luzin, „se apără" de această descoperire, despre care mărturisește că nu a avut nici cea mai vagă intuiție, arătînd în același timp valoarea și frumusețea incontestabilă a acestei teorii.

Deși, după cum vom vedea, putem ajunge la noțiunea de mulțime analitică fără a cunoaște noțiunea de mulțime boreliană, este totuși preferabil să amintim mai întîi definiția și unele proprietăți ale mulțimilor boreliene.

Fie o mulțime X. Toate mulțimile considerate mai jos vor fi părți ale lui X.

O familie S nevidă de mulțimi formează un corp, dacă operațiile de reuniune cel mult numărabilă și de diferență, aplicate mulțimilor din S, conduc tot la mulțimi din S.

Este ușor de văzut că fiind date mai multe corpuri de mulțimi, mulțimile care sînt comune tuturor acestor corpuri formează de asemenea un corp (chiar dacă familia de corpuri este infinită).

Fie acum o familie A de părți ale lui X. Să considerăm totalitatea corpurilor care conțin mulțimile familiei A. Există cel puțin un astfel de corp, de pildă cel format din toate părțile lui X. Intersecția acestor corpuri este tot un corp care conține pe A și este „cel mai mic" corp care conține mulțimile familiei A. Acest corp se numește corpul generat de A și se notează cu $S(A)$.

Să presupunem acum că X este spațiul euclidian cu n dimensiuni. Fie A familia tuturor intervalelor deschise n-dimensionale. Este ușor de văzut că A nu formează un corp. Corpul $S(A)$ se numește corpul mulțimilor boreliene din spațiul euclidian n-dimensional. Importanța acestui corp decurge din faptul că cele mai multe probleme care se pun în teoria clasică a funcțiilor de variabilă reală sau a celor de variabilă complexă, conduc la mulțimi boreliene. Luzin a comparat rolul mulțimilor boreliene în raport cu mulțimile în genere, cu acela al numerelor raționale în raport cu numerele reale. Intervalele sînt cele mai simple mulțimi boreliene. Importanța lor în analiză este binecunoscută. De cele mai multe ori, proprietățile unei funcții se studiază într-un interval. După intervale, cele mai simple mulțimi boreliene sînt mulțimile deschise și cele închise. O mulțime deschisă este o mulțime care, o dată cu un punct al ei, conține o întreagă sferă centrală în acest punct (sfera va trebui să aibă atîtea

dimensiuni, cîte dimensiuni are spațiul euclidian în care lucrăm). Se arată că orice mulțime deschisă este o reuniune numărabilă de intervale. Prin mulțime închisă se înțelege o mulțime a cărei complementară este deschisă. Mulțimea punctelor în care o funcție continuă întrece o valoare dată este totdeauna deschisă. Mulțimea punctelor în care o funcție continuă egalează o valoare dată este totdeauna închisă. Mulțimea punctelor de analiticitate ale unei funcții reale, de variabilă reală, sau ale unei funcții complexe de variabilă complexă este totdeauna deschisă.

Mulțimile boreliene care urmează în ordinea complicației sînt reuniunile numerabile de mulțimi închise (o astfel de reuniune se spune că este o mulțime de tipul F_σ) și intersecțiile numărabile de mulțimi deschise (o astfel de intersecție se spune că este o mulțime de tipul G_δ). Complementara unei mulțimi de tipul F_σ este o mulțime de tipul G_δ și reciproc. Fiind dată o funcție reală, arbitrară, de variabile reale, mulțimea punctelor ei de continuitate este de tipul G_δ; mulțimea punctelor de discontinuitate este de tipul F_σ. Mai mult, orice mulțime de tipul G_δ este mulțimea punctelor de continuitate ale unei anumite funcții reale.

Mulțimea numerelor raționale este de tipul F_σ, dar nu este de tipul G_δ. Mulțimile care sînt în același timp de tip F_σ și de tip G_δ au o mare importanță în teoria funcțiilor reale. Astfel, matematicianul sovietic A. S. Kronrod a arătat că pentru ca o mulțime liniară să fie atît de tip F_σ, cît și de tip G_δ, este necesar și suficient ca ea să fie mulțimea punctelor de discontinuitate ale unei funcții reale de o variabilă reală, derivabilă în orice punct în care este continuă,

O intersecție numărabilă de mulțimi de tipul F_σ constituie o mulțime de tipul $G_{\sigma\delta}$. O reuniune numărabilă de mulțimi de tipul G_δ constituie un $G_{\delta\sigma}$. Fiind dat un șir de funcții reale, continue, de variabilă reală, mulțimea punctelor de convergență ale acestui șir formează un $F_{\sigma\delta}$; reciproc, orice $F_{\sigma\delta}$ este mulțimea punctelor de convergență ale unui anume șir de funcții continue.

Continuînd astfel, prin aplicarea operațiilor de reuniune numărabilă și trecere la complementară, se obțin mulțimi

boreliene din ce în ce mai complicate. Mulțimile boreliene care apar cel mai des în analiza clasică sînt cele închise, cele deschise, cele de tip F_σ și cele de tip G_δ. S-ar putea face o listă foarte bogată de proprietăți care conduc la astfel de tipuri boreliene. Pe măsură ce studiul funcțiilor a progresat, s-au întîlnit din ce în ce mai des și tipuri boreliene mai complicate, mai cu seamă în spațiul euclidian unidimensional. Mulțimile boreliene de numere reale au unele proprietăți remarcabile. Astfel, fiind dată o funcție f reală, univalentă (orice valoare este luată cel mult într-un punct) de variabilă reală, care este limita unui șir convergent de funcții continue, nu numai că mulțimea valorilor lui f este boreliană, dar reciproc, fiind dată o mulțime boreliană de numere reale, există o funcție f, reală, univalentă, limită a unui șir convergent de funcții continue, astfel încît mulțimea valorilor lui f este însăși mulțimea boreliană dată. Această proprietate arată că orice mulțime boreliană, oricît de complicată ar fi ea, poate să apară în studiul unor funcții uzuale: funcțiile limită de funcții continue. Cele mai multe funcții uzuale în analiză se obțin ca limite de șiruri convergente de funcții continue. Așa sînt, de pildă, funcțiile derivate, funcțiile cu variație mărginită, funcțiile superior (sau inferior) semicontinue, funcțiile continue la stînga (sau la dreapta) în fiecare punct etc. Este deci clar că pentru teoria funcțiilor uzuale, este necesar să se întreprindă studiul proprietăților care aparțin tuturor mulțimilor boreliene și nu numai tipurilor boreliene mai simple.

Acest studiu a fost făcut și a dezvăluit proprietăți remarcabile. Astfel, orice mulțime boreliană este o imagine continuă a mulțimii numerelor iraționale. Orice imagine topologică (adică biunivocă și bicontinuă) a unei mulțimi boreliene este o mulțime boreliană. N. Luzin a pus unui elev al său, P. S. Aleksandrov (ulterior, șeful școlii sovietice de topologie), problema puterii mulțimilor boreliene. P. S. Aleksandrov a reușit să demonstreze, în 1916 (concomitent cu Hausdorff), că orice mulțime boreliană, nenumărabilă, conține o mulțime perfectă nevidă. Rezultatul lui P. S. Aleksandrov este de o mare importanță. Într-adevăr, se știa dinainte că orice mulțime perfectă

nevidă este de puterea continuului. Din teorema lui P. S. Aleksandrov rezultă deci că orice mulțime boreliană nenumărabilă este de puterea continuului. În felul acesta, mulțimile boreliene „realizează ipoteza continuului". (Ipoteza continuului afirmă că orice parte nenumărabilă a mulțimii numerelor reale este de puterea continuului, adică poate fi pusă în corespondență biunivocă cu mulțimea tuturor numerelor reale.) K. Gödel a demonstrat că ipoteza continuului nu vine în contradicție cu sistemul de axiome al teoriei mulțimilor, dacă acest sistem nu este contradictoriu în sine.

În 1963, Paul Cohen a coompletat rezultatul lui Gödel, demonstrînd că nici negația ipotezei continuului nu vine în contradicție cu sistemul de axiome al teoriei mulțimilor, dacă acest sistem nu este contradictoriu. Rezultă că ipoteza continuului constituie o propoziție independentă, așa cum de altfel am semnalat și în capitolul precedent.

Paralel cu mulțimile boreliene, se introduc funcțiile boreliene; o funcție reală f de o variabilă reală se numește boreliană, dacă mulțimea $\{x;\ f(x) > \alpha\}$ este boreliană, oricare ar fi α. Lebesgue a arătat că funcțiile boreliene coincid cu funcțiile din așa-numita clasificație a lui Baire, prezentată succint în capitolul. Funcțiile continue constituie funcțiile de clasa zero. Orice funcție discontinuă care este limita unui șir convergent de funcții continue este o funcție de primă clasă. Orice funcție care nu este continuă sau de prima clasă, dar care este limita unui șir convergent de funcții continue sau de prima clasă, este o funcție de-a doua clasă și așa mai departe. Dacă o funcție nu este de nici o clasă finită, dar se poate obține ca limită a unui șir convergent de funcții de clasă finită, atunci se spune că această funcție este de clasă ω (numită și prima clasă transfinită). Cu ajutorul funcțiilor de clasă ω se definesc acum funcțiile de clasă $\omega + 1$, apoi cele de clasă $\omega + 2$ ș.a.m.d. Principiile după care se definesc aceste clase sînt următoarele: mulțimea claselor lui Baire este nenumărabilă; fiecare clasă este precedată de un număr finit sau o infinitate numărabilă de clase; după orice mulțime numărabilă de clase urmează o nouă clasă. Clasa funcțiilor boreliene este cea mai mică clasă

de funcții care conține toate polinoamele și astfel încît, pentru orice șir convergent de funcții din clasă, limita șirului aparține clasei.

Cele mai multe funcții discontinue care intervin în analiză sînt de prima clasă Baire. Totuși, încă în secolul trecut, Dirichlet a fost condus în mod natural să considere, de pildă, funcția $f(x) = \lim\limits_{m\to\infty} (\lim\limits_{n\to\infty} \cos (m!\, \pi x)^{2n})$, funcție care este de a doua clasă Baire (această funcție este egală cu 1 pentru x rațional și cu 0 pentru x irațional).

CÎTEVA EXEMPLE „RARE"

Există mulțimi neboreliene? Există funcții de clasă Baire oricît de mare? Există funcții care nu aparțin clasificației lui Baire? Iată întrebări care s-au pus din primul moment al apariției teoriilor lui Borel și Baire. Răspunsul, la toate întrebările puse, este afirmativ și se obține fără mari dificultăți. Se poate arăta că familia tuturor mulțimilor boreliene este de puterea continuului. Pe de altă parte, se poate arăta că familia tuturor mulțimilor din spațiul euclidian este de putere mai mare decît a continuului. În felul acesta, rezultă existența mulțimilor neboreliene. Mai mult, se arată că orice mulțime de puterea continuului conține o submulțime neboreliană.

Acum, ținînd seama de teorema lui Lebesgue asupra identității dintre funcțiile boreliene și funcțiile din clasificația lui Baire, rezultă ușor existența funcțiilor reale, de variabilă reală, care nu aparțin clasificației lui Baire. Este suficient să considerăm o mulțime E neboreliană și să punem $f(x) = 1$, dacă $x \in E$ și $f(x) = 0$ dacă $x \notin E$.

Demonstrația de mai sus a nemulțumit pe unii matematicieni, din cauza caracterului ei neefectiv. Într-adevăr, demonstrația dată nu indică un exemplu de mulțime sau de funcție neboreliană, ci ne face cunoștință cu existența mulțimilor și a funcțiilor neboreliene pe o cale oarecum indirectă.

Nu este deci lipsit de interes un exemplu de funcție discontinuă, care nu este nici de prima, nici de-a doua clasă Baire.

Raționamentul pe care-l prezentăm se aplică și pentru a da un exemplu de funcție care nu aparține nici unei clase inferioare unei clase date.

Să punem

$$P_{\alpha,\beta}(x) = \sum_{\gamma=0}^{\gamma=\alpha+\beta} C_{\alpha,\beta,\gamma}\, x^{\gamma}.$$

Orice serie dublă de polinoame poate fi scrisă sub forma

$$\sum_{\alpha=1}^{\infty}\left[\sum_{\beta=1}^{\infty} P_{\alpha,\beta}(x)\right]. \qquad (1)$$

Dacă pe un anumit interval seria (1) converge, atunci ea definește pe acest interval o funcție de clasă 0, 1 sau 2. Reciproc, orice funcție de clasă 0,1 sau 2 poate fi definită în acest fel (deoarece se știe că orice funcție continuă este limită uniformă de polinoame) printr-o alegere convenabilă a coeficienților $C_{\alpha,\beta,\gamma}$. Fiecărui sistem de constante $C_{\alpha,\beta,\gamma}$ să-i asociem un număr real u cuprins între 0 și 1 în așa fel, încît corespondența între aceste sisteme de constante și numerele din [0,1] să fie biunivocă.

Să definim acum pe (0, 1) o funcție reală f în felul următor: fie $0 \leqslant u \leqslant 1$. Fie $C_{\alpha,\beta,\gamma}$ sistemul de constante corespunzător lui u. Fie (1) seria formată cu acest sistem de constante. Dacă seria (1) astfel formată, diverge sau converge către un număr diferit de zero, pentru $x = u$ punem $f(u) = 0$; dacă, dimpotrivă, seria (1) are ca sumă pe zero pentru $x = u$, punem $f(u) = 1$. Funcția f astfel definită nu coincide cu nici o funcție de clasă 0, 1 sau 2, deoarece o astfel de funcție, definită pe [0, 1], este reprezentabilă printr-o serie (1) convergentă pentru toate valorile lui x cuprinse între 0 și 1. Să presupunem, pentru o clipă, că f se poate reprezenta printr-o astfel de serie. Fie u numărul din [0,1] care corespunde sistemului $C_{\alpha,\beta,\gamma}$ de coeficienți din (1); pentru $x = u$ seria converge, deoarece u este cuprins între 0 și 1; dacă suma seriei este diferită de zero, trebuie să avem $f(u) = 0$, iar dacă suma seriei este egală cu zero, trebuie să avem $f(u) = 1$. Și într-un caz și în celălalt obținem o contradicție.

Din cele arătate rezultă că f nu este de clasă 0, 1 sau 2. Dar nu rezultă, cel puțin explicit, că f aparține clasificației lui Baire sau, în caz afirmativ, cărei clase îi aparține.

Se poate obține, folosind anumite modificări și completări ale raționamentului de mai sus, o funcție de clasă 3. Lebesgue a demonstrat că nici o clasă (finită sau transfinită) din clasificația lui Baire nu este vidă. Totodată, el a arătat că se pot defini, fără a utiliza axioma lui Zermelo, funcții care nu aparțin clasificației lui Baire.

Nu este locul să insistăm aici asupra unei chestiuni care a dat naștere la multe discuții. Existența unei ființe matematice trebuie considerată întotdeauna în raport cu metoda prin care ea se stabilește. Demonstrațiile de existență ale lui Lebesgue, despre care am vorbit mai sus, utilizează fie diagonala lui Cantor (o generalizare a metodei prin care se arată că mulțimea numerelor reale este nenumărabilă) fie totalitatea numerelor transfinite, de prima și a doua clasă. Numerele transfinite de prima și a doua clasă transfinită pot fi definite ca indicii de ordine ai claselor din clasificația lui Baire. Totalitatea lor este, evident, nenumărabilă. Aceste metode sînt socotite mai efective decît, de pildă, utilizarea axiomei alegerii a lui Zermelo, utilizare care este considerată cazul cel mai tipic de demonstrație neefectivă. Dar ele nu pot fi considerate în nici un caz niște raționamente constructive. Baire a dat un exemplu aritmetic (deci constructiv) de funcție de a treia clasă. Să expunem acest exemplu. Fie dezvoltarea în fracție continuă a fiecărui număr x^*, cuprins între 0, 1; $x = \frac{1|}{|\alpha_1} + \frac{1|}{|\alpha_2} + \frac{1|}{|\alpha_3} + \cdots$. Să definim pe $[0, 1]$ o funcție f în felul următor: $f(x) = 1$, dacă șirul α_n tinde către $+\infty$, $f(x) = 0$ în cazul contrar. Se poate arăta că această funcție este de a treia clasă Baire.

Ludmila Keldîș a reușit să construiască un exemplu de funcție de a patra clasă Baire. Iată acest exemplu. Fie, din nou, dezvoltarea în fracție continuă a fiecărui x cuprins între 0 și 1. Dacă în șirul $\alpha_1, \alpha_2, \alpha_3 \ldots$, format cu cîturile necom-

---

* Este preferabil, în cele două exemple care urmează, ca x să fie irațional, pentru ca dezvoltarea lui x în fracție continuă să fie totdeauna esențial infinită.

plete ale dezvoltării lui x, există o infinitate de termeni distincți, care se repetă, fiecare, de o infinitate de ori, atunci punem $f(x) = 1$. În cazul contrar, punem $f(x) = 0$.

Ulterior, Ludmila Keldîș a obținut un rezultat și mai puternic. Ea a dat un exemplu pur aritmetic de funcție de a n-a clasă Baire, n fiind un număr natural dat.

Dacă insistăm asupra diverselor tipuri de existență în matematică, o facem pentru că distincția între aceste tipuri are o mare importanță în teoria mulțimilor analitice și proiective. Lebesgue a obținut, fără să-și dea seama, o mulțime analitică neborelieană. El nu i-a acordat atenția cuvenită, pentru că ea juca un rol accesoriu în problema care-l preocupa, aceea a definirii unei funcții care scapă oricărei reprezentări analitice (deci care este în afara clasificației lui Baire). Lebesgue a utilizat în mod esențial, în demonstrația existenței unei astfel de funcții, totalitatea numerelor transfinite de prima și a doua clasă. Vom vedea mai tîrziu că în acea clasificare a mulțimilor care ia ca punct de plecare mulțimile analitice, existența mulțimilor de diverse clase se demonstrează prin procedeul diagonal al lui Cantor.

Este un mare merit al lui Luzin de a fi observat că, după cum se exprimă el însuși, în construcția mulțimii definite de Lebesgue „se află conținută, ca într-un germene, întreaga teorie a mulțimilor analitice". De această mulțime se leagă așa-numita teorie geometrică a mulțimilor analitice, teorie construită în întregime de Luzin. Ceva mai tîrziu, după ce vom prezenta punctul de vedere al lui Suslin, vom da unele elemente ale teoriei geometrice a lui Luzin și vom defini operația de trecere prin sită (sau de ciuruire), care joacă un rol principal în această teorie.

MULȚIMI ANALITICE ȘI OPERAȚIA LUI SUSLIN

Am anunțat că introducerea noțiunii de mulțime analitică are ca punct de plecare clasa mulțimilor boreliene și utilizează, ca operație fundamentală, proiecția.

Cele mai simple mulțimi boreliene din plan sînt dreptunghiurile cu laturile paralele cu axele. Partea comună a unor

astfel de dreptunghiuri este sau tot un astfel de dreptunghi, sau mulțimea vidă. Proiecția pe axa absciselor a unui astfel de dreptunghi este un interval unidimensional. Rezultă că cele mai simple mulțimi boreliene plane—intervalele bidimensionale — nu sînt în stare să furnizeze nimic nou atunci cînd sînt supuse operațiilor de intersecție și de proiecție.

Să trecem acum la clasa următoare de mulțimi boreliene, mulțimile deschise. O mulțime plană deschisă este o reuniune cel mult numărabilă de dreptunghiuri de tipul precedent. Însă este ușor de văzut că proiecția unei reuniuni de mulțimi este egală cu reuniunea proiecțiilor acestor mulțimi. Rezultă astfel că proiecția pe axa absciselor a unei mulțimi plane deschise este o mulțime deschisă liniară. Nici de astă dată nu am obținut ceva nou, nu am ieșit din clasa boreliană de la care am pornit.

Să trecem acum la mulțimile de tip G_δ. Fie A o astfel de mulțime. A este o intersecție cel mult numărabilă de mulțimi deschise. Aici trebuie să fim mai atenți, deoarece am văzut mai sus că proiecția unei intersecții nu este, în general, egală cu intersecția proiecțiilor.

Avem:

$$A = \bigcap_{m=1}^{\infty} \bigcup_{n=1}^{\infty} D_n^m = (D_1^1 \cup D_2^1 \cup D_3^1 \cup \cdots) \cap (D_1^2 \cup D_2^2 \cup \ldots) \cap$$

$$\cap (D_1^3 \cup D_2^3 \cup \cdots) \cap \cdots$$

sau

$$A = \bigcup_{(n_1, n_2, n_3, \ldots)} D_{n_1}^1 D_{n_2}^2 D_{n_3}^3 \cdots *). \tag{2}$$

Reuniunea care furnizează pe A parcurge toate șirurile posibile de numere naturale, deci este o reuniune nenumărabilă. $D_{n_i}^m$ sînt dreptunghiuri cu laturi paralele cu axele. Avem:

$$pr\, A = \bigcup_{(n_1, n_2, n_3, \ldots)} pr(D_{n_1}^1 D_{n_2}^2 D_{n_3}^3 \ldots), \tag{3}$$

unde prin pr am notat, și vom nota și de acum înainte, operația de proiecție pe axa absciselor. Acum vom profita

---

* Uneori vom nota intersecția unor mulțimi prin simpla juxtapunere a acestor mulțimi.

de o observație simplă, în legătură cu proiecția unei intersecții de mulțimi: fiind dat un șir descrescător de mulțimi plane, închise și mărginite, proiecția intersecției mulțimilor din șir este egală cu intersecția proiecțiilor acestor mulțimi. Dreptunghiurile care figurează în (3) sînt mulțimi închise și mărginite, însă șirul $D^1_{n_1}$, $D^2_{n_2}$, $D^3_{n_3}$, ... nu este, în general, un șir descrescător. Va trebui deci să înlocuim acest șir printr-unul descrescător, care însă să aibă aceeași intersecție ca și primul. Iată cum se poate face aceasta: punem

$$D^1_{n_1} \cap D^2_{n_2} \cap D^3_{n_3} \cap \ldots \cap D^k_{n_k} = D_{n_1,\, n_2,\, \ldots n_k}.$$

Este vizibil că

$$D^1_{n_1} \cap D^2_{n_2} \cap D^3_{n_3} \cap \ldots = D_{n_1} \cap D_{n_1 n_2} \cap D_{n_1 n_2 n_3} \cap \ldots . \quad (4)$$

În același timp avem

$$D_{n_1} \supset D_{n_1 n_2} \supset D_{n_1 n_2 n_3} \supset \cdots ,$$

deci

$$pr(D_{n_1} \cap D_{n_1 n_2} \cap D_{n_1 n_2 n_3} \cap \ldots) = prD_{n_1} prD_{n_1 n_2} prD_{n_1 n_2 n_3} \ldots \quad (5)$$

Să punem

$$prD_{n_1 n_2 \ldots n_k} = \delta_{n_1 n_2 \cdots n_k}. \quad (6)$$

$\delta_{n_1 n_2 \cdots n_k}$, ca proiecție a unui dreptunghi, este un interval unidimensional sau mulțimea vidă (dacă $D_{n_1 n_2 \cdots n_k}$ este mulțimea vidă).
Ținînd seama de (3), (4) (5) și (6), obținem

$$pr\, A = \bigcup_{(n_1,\, n_2,\, n_3 \ldots)} \delta_{n_1}\, \delta_{n_1 n_2}\, \delta_{n_1 n_2 n_3} \cdots . \quad (7)$$

Faptul remarcabil este că nu numai mulțimile plane de tip G_δ, dar orice mulțime boreliană plană furnizează, prin proiecție, o mulțime de tipul membrului doi din (7). Este astfel de înțeles de ce tocmai expresia (7) a atras atenția principală a lui Mihail Suslin. Suslin numește sistem determinant de mulțimi o familie de mulțimi $\{A_{n_1 n_2 \ldots n_k}\}$ cu proprietatea că fiecărui sistem finit n_1, n_2, ... n_k de

numere naturale îi corespunde o mulțime din familie, pe care o notăm cu $A_{n_1 n_2 \ldots n_k}$. Numim nucleu al sistemului determinant o mulțime pe care o notăm cu $N\{A_{n_1 n_2 \ldots n_k}\}$ și care are expresia

$$\bigcup_{(n_1, n_2, \ldots, n_k, \ldots)} A_{n_1} A_{n_1 n_2} A_{n_1 n_2 n_3} \ldots .$$

Se numește mulțime analitică liniară sau mulțime A nucleul oricărui sistem determinant format din intervale unidimensionale închise. Dacă un sistem determinant este format din intervale bidimensionale închise, nucleul său este o mulțime analitică plană. Ținînd seama că mulțimea tuturor șirurilor de numere naturale este de puterea continuului, rezultă că familia mulțimilor analitice este de puterea continuului. Să mai observăm că din faptul că un interval este o mulțime analitică și că reuniunea cel mult numărabilă și intersecția cel mult numărabilă de mulțimi analitice conduc tot la mulțimi analitice, rezultă următorul fapt extrem de important: orice mulțime boreliană este analitică. Din definiția pusă și din dezvoltările efectuate rezultă că proiecția oricărei mulțimi boreliene este o mulțime analitică. Mai mult decît atît, avem în aceasta o proprietate caracteristică, o nouă definiție a mulțimilor analitice. În adevăr, fie M o mulțime analitică liniară, Fie $\{\delta_{n_1 n_2 \ldots n_k}\}$ sistemul determinant format din intervale închise, care admite ca nucleu pe M, deci pentru care

$$M = \bigcup_{(n, n_2, \ldots, n_k \ldots)} \delta_{n_1} \delta_{n_1 n_2} \delta_{n_1 n_2 n_3} \ldots . \tag{8}$$

Înlocuim fiecare interval $\delta_{n_1 n_2 \ldots n_k}$ prin intersecția $\delta_{n_1} \delta_{n_1 n_2} \ldots \delta_{n_1 n_2 \ldots n_k}$. Aceasta nu schimbă intersecția $\delta_{n_1} \delta_{n_1 n_2} \delta_{n_1 n_2 n_3} \cdots \delta_{n_1 \ldots n_k} \ldots$ care figurează în (8), transformînd-o în același timp într-o intersecție a unui șir descrescător de mulțimi. Pentru a nu mai complica însă scrierea, admitem că

$$\delta_{n_1} \supset \delta_{n_1 n_2} \supset \delta_{n_1 n_2 n_3} \supset \cdots . \tag{9}$$

Să asociem fiecărui δ_k un dreptunghi D_k, a cărui proiecție pe axa absciselor este δ_k și astfel încît dacă $m \neq n$, D_m nu

are puncte comune cu D_n. În interiorul fiecărui dreptunghi D_{n_1}, să construim acum dreptunghiuri D_{n_1, n_2} ($n_2 = 1, 2, 3 \ldots$), ale căror proiecții pe axa absciselor sînt intervalele $\delta_{n_1 n_2}$. Aceasta este posibil, deoarece $\delta_{n_1} \supset \delta_{n_1 n_2}$ (a se vedea (9)). Dreptunghiurile $D_{n_1 n_2}$ vor fi construite în așa fel, încît dacă $p \neq q$, atunci $D_{n_1 p}$ nu are puncte comune cu $D_{n_1 q}$. Continuînd astfel, obținem dreptunghiuri $D_{n_1 n_2 \ldots n_k}$, ale căror proiecții pe axa absciselor sînt tocmai intervalele $\delta_{n_1 n_2 \ldots n_k}$, astfel încît fiecare dreptunghi $D_{n_1 n_2 \ldots n_k}$ este conținut în dreptunghiul $D_{n_1 n_2 \ldots n_{k-1}}$, iar dreptunghiurile $D_{n_1 n_2 \ldots n_{kp}}$ și $D_{n_1 n_2 \ldots n_{kq}}$ nu au puncte comune dacă $p \neq q$.

Fie acum E_k reuniunea tuturor dreptunghiurilor $D_{n_1 n_2 \ldots n_k}$, formate cu k indici. E_k este o reuniune numărabilă de mulțimi boreliene, deci este ea însăși boreliană. Mulțimea $A = \bigcap_{k=1}^{\infty} E_k$ este deci de asemenea boreliană. Rămîne să observăm acum că $pr\ A = M$, M fiind mulțimea analitică dată de (8). Prin aceasta am demonstrat că mulțimile analitice liniare coincid cu proiecțiile mulțimilor boreliene din plan.

W. Sierpiński a arătat că nu este necesar, pentru a obține totalitatea mulțimilor analitice liniare, să considerăm proiecțiile mulțimilor boreliene de toate tipurile posibile; el a arătat că orice mulțime analitică liniară este proiecția unei mulțimi plane, de tipul borelian G_δ.

LEBESGUE, AUTOR INVOLUNTAR AL PRIMULUI EXEMPLU DE MULȚIME ANALITICĂ NEBORELIANĂ

Am văzut că descoperirea lui Suslin își are originea într-o greșeală din memoriul citat al lui Lebesgue. Dar teoria mulțimilor analitice mai are și alte puncte de contact cu acest memoriu. Unul dintre cele mai importante a fost sesizat de Luzin. Iată despre ce este vorba.

La sfîrșitul memoriului său, Lebesgue efectuează următoarea construcție: fie $r_1, r_2, \ldots, r_i, \ldots$ un șir format cu toate numerele raționale din $[0, 1]$. Fie, pentru un x dat

din [0, 1], dezvoltarea diadică corespunzătoare (avînd grijă ca cifra 1 să figureze de o infinitate de ori).

$$x = \frac{\theta_1}{2} + \frac{\theta_2}{2^2} + \cdots + \frac{\theta_i}{2^i} + \cdots .$$

Să suprimăm din șirul $\{r_i\}$ pe toți acei r_i pentru care $\theta_i = 0$. Fie $r'_1\, r'_2 \cdots, r'_i, \cdots$ șirul rămas. Să considerăm acum mulțimea E a numerelor x din [0, 1] pentru care punctele corespunzătoare valorilor r'_i formează o mulțime bine ordonată în raport cu relația $<$ (o mulțime de numere reale este bine ordonată, dacă orice parte a ei conține un cel mai mic element). Luzin arată că complementara A a mulțimii E este o mulțime analitică neboreliană.

Acesta este primul exemplu, în ordine cronologică, de mulțime analitică neboreliană, exemplu care a stat la baza cercetărilor lui Luzin și Suslin.

Pentru a înțelege valoarea exemplului lui Lebesgue, trebuie să nu uităm că familia mulțimilor analitice este de puterea continuului și, deci, că un exemplu de mulțime analitică neboreliană nu se poate obține cu ajutorul axiomei lui Zermelo, deoarece această axiomă, conducînd la c^c alegeri posibile, ar furniza o clasă de mulțimi de putere superioară continuului.

Pentru a demonstra analiticitatea mulțimii A, Luzin a folosit următoarea variantă a noțiunii de mulțime analitică (se demonstrează echivalența ei cu definiția dată mai sus): „O mulțime liniară este analitică, dacă ea este mulțimea valorilor unei funcții reale de variabilă reală ale cărei puncte de discontinuitate formează o mulțime cel mult numărabilă".

Acum putem înțelege sensul denumirii „analitică". Orice funcție de tipul de mai sus este limita unui șir convergent de funcții continue. De aici se deduce că o astfel de funcție este suma unei serii convergente de polinoame cu coeficienți întregi, deci o astfel de funcție este definită cu ajutorul unei expresii analitice.

Mulțimea A a sugerat lui Luzin introducerea operației de *ciuruire* sau de trecere prin sită. Fie două axe perpendiculare și fie pătratul cu laturile date de ecuațiile $x = 0$, $y = 0$, $x = 1$, $y = 1$. Să considerăm pe latura $x = 0$ a acestui pătrat toate punctele raționale r_1, r_2, $\cdots$ r_n, $\cdots$ (cuprinse între 0 și 1). Să ducem prin r_n paralela la axa $y = 0$ și să împărțim porțiunea din paralelă, cuprinsă în pătratul de mai sus, în 2^n segmente egale. Să reținem dintre aceste segmente numai pe cele de rang par (rangul fiind socotit de la stînga spre dreapta). Efectuînd această operație pentru fiecare n, obținem o familie numărabilă de segmente paralele cu axa $y = 0$ și formînd o mulțime G densă în pătratul de mai sus. Este natural să numim sită reuniunea acestor segmente. Într-adevăr, orice dreaptă $x = x_0$ $(0 < x_0 \leqslant 1)$ întîlnește în mulțimea G o mulțime de puncte de ordonate raționale, mulțime care se proiectează pe axa $x = 0$ după o parte R_{x_0} a mulțimii $\{r_1, r_2, \ldots, r_n, \ldots\}$. Deci mulțimea G introduce, pentru fiecare paralelă la axa $x = 0$, o selecție în mulțimea $\{r_1, r_2, \ldots, r_n, \cdots\}$, reținînd doar o parte a acestei mulțimi. Este deci o veritabilă sită.

Dar lucrurile nu se opresc aici. Observăm că un punct r_n de pe $x = 0$ aparține proiecției R_x, dacă și numai dacă avem $\theta_n = 1$ (θ_n fiind a n-a cifră în dezvoltarea diadică a lui x_0). De aici rezultă că mulțimea A se caracterizează prin proprietatea că orice dreaptă $x = x_0$, cu $x_0 \in A$ (și numai o astfel de dreaptă), întîlnește în sita G o mulțime care nu este bine ordonată, deci pentru care R_{x_0} nu este bine ordonată (în raport cu ordinea în care punctele se întîlnesc de jos în sus). Spunem că mulțimea A este obținută prin ciuruire, cu ajutorul sitei G, prin mijlocirea proprietății de a nu fi bine ordonată.

Fie acum, în general, o mulțime M densă în plan. Fie π o proprietate care are sens pentru mulțimile liniare. Fie $\Lambda_\pi(M)$ mulțimea tuturor numerelor reale x_0, astfel încît mulțimea de intersecție a dreptei $x = x_0$ cu M are proprietatea π. Spunem că $\Lambda_\pi(M)$ se obține ciuruind mulțimea plană M cu ajutorul proprietății π.

Iată acum cîteva rezultate remarcabile legate de operația de ciuruire (trecere prin sită). C. Kuratovski și E. Szpilrajn-Marczewski au arătat că ori de cîte ori M este închisă, iar π este proprietatea de a conține cel puțin un punct de ordonată irațională, $\Lambda_\pi(M)$ este o mulțime analitică. S. Mazurkiewicz și W. Sierpiński au arătat că dacă M este închisă, iar π este proprietatea de a fi nenumărabilă, $\Lambda_\pi(M)$ este analitică. În sfîrșit, N. Luzin și W. Sierpiński au arătat că dacă M este închisă, iar π proprietatea de a nu fi bine ordonată după mărimea ordonatelor, $\Lambda_\pi(M)$ este de asemenea analitică.

MULȚIMILE ANALITICE ȘI MULȚIMEA NUMERELOR IRAȚIONALE

Am semnalat mai sus că orice mulțime analitică poate fi obținută cu ajutorul unei funcții care are o mulțime cel mult numărabilă de puncte de discontinuitate. Acesta nu este singurul mod de a obține o mulțime analitică drept mulțime a valorilor unei anumite funcții. Iată cum se poate ajunge la un alt rezultat de acest tip: se poate arăta că orice mulțime analitică poate fi obținută ca nucleu al unui sistem determinant $\{\delta_{n_1 n_2, \ldots, n_k}\}$ pentru care intervalele închise $\delta_{n_1, n_2, \ldots, n_k}$ sînt nevide și verifică relația (9), iar lungimea intervalului $\delta_{n_1 n_2 \ldots n_k}$ este inferioară lui $\frac{1}{k}$. În acest fel, intersecția $\delta_{n_1} \delta_{n_1 n_2} \delta_{n_1 n_2 n_3} \ldots$, care apare în expresia (8) a unei mulțimi analitice, se reduce la un singur punct. Dar șirul $n_1, n_2, n_3, \ldots$ definește un număr irațional, și anume (se știe că un număr real admite o dezvoltare în fracție continuă infinită, dacă și numai dacă el este irațional):

$$x = \cfrac{1}{n_1 + \cfrac{1}{n_2 + \cfrac{1}{n_3 + \ldots}}}$$

Rezultă că mulțimea analitică dată coincide cu mulțimea valorilor funcției $f(x)$, definite pe mulțimea I a numerelor

iraționale cuprinse între 0 și 1, prin condiția ca $f(x)$ să fie punctul comun intervalelor δ_{n_1}, $\delta_{n_1 n_2}$, $\delta_{n_1 n_2 n_3}$, ... Această funcție este continuă pe și prin mulțimea I. Într-adevăr, fiind date două numere $x \in I$, $x' \in I$, $x = \frac{1|}{|n_1} + \frac{1|}{|n_2} + \frac{1|}{|n_3} + \cdots$ și $x' = \frac{1|}{|n_1'} + \frac{1|}{|n_2'} + \cdots$, avem $n_i = n_i'$, $1 \leqslant i \leqslant k$, pentru un k cu atît mai mare cu cît x și x' sînt mai apropiate. Însă în aceste condiții, $f(x)$ și $f(x')$ aparțin unui același interval $\delta_{n_1 n_2 \ldots n_k}$, unde k este cu atît mai mare, cu cît x și x' sînt mai apropiate. Amintindu-ne că lungimea lui $\delta_{n_1 n_2 \ldots n_k}$ este inferioară lui $\frac{1}{k}$, rezultă continuitatea funcției $f(x)$. Am stabilit deci că orice mulțime analitică liniară este imagine continuă a mulțimii I. Însă mulțimea I este la rîndul ei o imagine continuă a mulțimii tuturor numerelor iraționale. În concluzie: orice mulțime analitică liniară este o imagine continuă a mulțimii tuturor numerelor iraționale. Acest rezultat nu este de natură să ne surprindă, deoarece el este cunoscut dinainte (a se vedea prima parte a capitolului) pentru o clasă vastă de mulțimi analitice, aceea a mulțimilor boreliene. Vom arăta însă — și aceasta nu mai este valabil pentru mulțimile boreliene — că și reciproca teoremei de mai sus este adevărată. Fie, într-adevăr, $f(x)$ o funcție continuă pe și prin mulțimea R a tuturor numerelor iraționale. Fie G graficul lui $f(x)$, adică mulțimea plană a punctelor de coordonate $(x, f(x))$, unde $x \in R$. Mulțimea A a valorilor lui $f(x)$ este proiecția, pe axa ordonatelor, a mulțimii plane G. Este deci suficient, pentru a demonstra analiticitatea mulțimii A, să arătăm că G este o mulțime boreliană (deoarece s-a arătat mai sus că mulțimile analitice liniare coincid cu proiecțiile mulțimilor boreliene din plan). Pentru aceasta, vom arăta că G este intersecția a două mulțimi boreliene. Fie $\bar{G}$ închiderea lui G, adică totalitatea punctelor lui G și a punctelor de acumulare ale lui G. Este ușor de văzut că $\bar{G}$ este închisă, deci boreliană. Fie Γ mulțimea tuturor punctelor din plan a căror abscisă este irațională. Γ este o mulțime de tip G_δ, deoarece complementara ei este reuniunea (numărabilă) a dreptelor $x = r$ (r rațional),

iar o dreaptă este o mulțime închisă. Din faptul că $f(x)$ este continuă pe și prin R, rezultă că $G = \bar{G} \cap \Gamma$, deci G este boreliană (G este de tip G_δ, deoarece se arată ușor că orice mulțime închisă este de tip G_δ). Am demonstrat deci că orice imagine continuă a mulțimii numerelor iraționale este o mulțime analitică.

Să mai semnalăm, fără a intra în detalii, că orice mulțime analitică liniară este mulțimea valorilor unei funcții superior semicontinue (adică pentru care $\lim_{x' \to x} f(x') \leqslant f(x)$, oricare ar fi x) și că valorile oricărei funcții din clasificația lui Baire formează o mulțime analitică. (Valorile unei funcții măsurabile nu formează, în general, nici măcar o mulțime măsurabilă.)

CUM RECUNOAȘTEM DACĂ O MULȚIME ANALITICĂ ESTE BORELIANĂ ?

Prima problemă mai importantă care s-a pus în teoria mulțimilor analitice a fost aceea a existenței unei mulțimi analitice neboreliene. Astfel, scopul principal al lui Suslin a fost găsirea unei mulțimi boreliene plane, a cărei proiecție pe o dreaptă este o mulțime liniară neboreliană. Am văzut mai sus că și Luzin s-a ocupat mult de această problemă. După rezolvarea ei, era important să se stabilească criterii pentru recunoașterea mulțimilor boreliene. Primul criteriu de acest tip a fost stabilit de Suslin, care a arătat că o condiție necesară și suficientă ca o mulțime să fie boreliană este ca atît ea, cît și complementara ei să fie analitice. Acest rezultat este foarte important, deoarece pune în evidență existența unei clase noi de mulțimi, clase care se situează în mod natural după clasa mulțimilor analitice. Studiul lor s-a dovedit a fi dificil și, după cum vom arăta, matematicienii nu au putut încă să rezolve principalele probleme legate de ele.

Ulterior s-au dat și alte criterii de recunoaștere a mulțimilor boreliene. Astfel, Luzin a arătat că o condiție necesară și suficientă ca o mulțime nenumărabilă (orice mulțime numărabilă este boreliană) să fie boreliană este ca ea să fie imaginea biunivocă și continuă într-un sens a mulțimii

numerelor iraționale. De aici rezultă că noțiunea de mulțime boreliană nu este numai un invariant topologic, ci și un invariant al unor transformări mai generale decît cele topologice, anume al transformărilor biunivoce și continue într-un sens.

PROPRIETĂȚI ALE MULȚIMILOR ANALITICE

O dată precizat locul mulțimilor boreliene în clasa mulțimilor analitice, o dată stabilit caracterul mai larg al acesteia din urmă, este natural să ne întrebăm care sînt proprietățile de structură ale mulțimilor analitice și comportarea lor față de diferite operații. Este ușor de văzut că o reuniune sau o intersecție numărabilă de mulțimi analitice conduce la o mulțime analitică.

O mulțime analitică nenumărabilă conține totdeauna o mulțime perfectă nevidă, deci are puterea continuului (pentru că o mulțime perfectă nevidă are puterea continuului). Orice mulțime analitică este măsurabilă în sens Lebesgue, dar există mulțimi măsurabile Lebesgue care nu sînt analitice.

O proprietate remarcabilă a mulțimilor analitice este dată de așa-numitul primul principiu al mulțimilor analitice, separabilitatea (B); fiind date două mulțimi analitice A și B, fără puncte comune, există două mulțimi boreliene A_1 și B_1, astfel încît $A \subset A_1$, $B \subset B_1$ și $A_1 \cap B_1 = 0$. Din această proprietate, găsită de Luzin, rezultă imediat jumătate din teorema de mai sus a lui Suslin. Într-adevăr, dacă o mulțime A și complementara ei B ar fi analitice, atunci am avea $A_1 = A$, $B_1 = B$, deci A ar fi boreliană.

Pe de altă parte, s-a putut da un exemplu de două mulțimi A și B complementare analitice fără puncte comune, pentru care nu există două mulțimi boreliene A_1 și B_1 fără puncte comune, astfel încît $A \subset A_1$, $B \subset B_1$.

Al doilea principiu al mulțimilor analitice, găsit tot de Luzin, este numit separabilitatea (CA). Pentru a-l putea enunța, să punem mai întîi o definiție: spunem că două mulțimi A și B, fără puncte comune, sînt separabile cu ajutorul a două complementare analitice, dacă există două

mulțimi A_1 și B_1 complementare analitice, astfel încît $A \subset A_1$, $B \subset B_1$, $A_1 \cap B_1 = 0$. Principiul al doilea se enunță astfel: dacă din două mulțimi analitice se suprimă partea lor comună, atunci cele două părți rămase sînt separabile (CA).

Există o complementară analitică nenumărabilă care nu conține nici o mulțime perfectă nevidă? Aceasta este una dintre problemele cele mai grele privitoare la complementarele analitice. Asupra caracterului acestei probleme vom reveni mai tîrziu, cînd vom discuta și despre alte probleme grele.

Să reținem deocamdată că o complementară analitică este totdeauna măsurabilă în sens Lebesgue, deoarece mulțimile analitice sînt măsurabile.

Am văzut mai sus că noțiunea de mulțime analitică este un invariant al transformărilor continue. Complementarele analitice nu au această proprietate. S-a putut însă arăta că noțiunea de complementară analitică este un invariant topologic.

În sfîrșit, o legătură între mulțimile analitice și complementarele lor este dată de următoarea teoremă: orice mulțime analitică este o imagine biunivocă și continuă a unei complementare analitice.

Nu vom mai insista asupra proprietăților mulțimilor analitice și ale complementarelor lor. Vrem acum să ilustrăm, prin exemple cît mai variate, felul în care intervine în matematică aceste mulțimi și mai ales acele dintre ele care nu sînt boreliene.

EXEMPLE DE MULȚIMI ANALITICE

Exemplul 1. Fie $f(x)$ o funcție reală, definită și continuă pe dreapta reală. Mulțimea valorilor luate de $f(x)$ în punctele ei de derivabilitate este analitică. Reciproc fiind dată o mulțime analitică liniară E, există o funcție reală $f(x)$, definită și continuă pe dreapta reală, astfel încît mulțimea valorilor luate de (fx) în punctele de derivabilitate coincide cu E.

Exemplul 2. Fie $f(x)$ o funcție reală, definită și continuă pe dreapta reală. Valorile pe care $f(x)$ le ia de o infinitate

nenumărabilă de ori formează o mulțime analitică. Reciproc, orice mulțime analitică liniară este mulțimea valorilor pe care o anumită funcție continuă le ia de o infinitate nenumărabilă de ori.

Exemplul 3. Valorile pe care o funcție reală boreliană, definită pe o mulțime boreliană de numere reale, le ia o singură dată formează o complementară analitică. Orice complementară analitică liniară este mulțimea valorilor pe care o anumită funcție reală, continuă, definită pe o mulțime închisă, le ia o singură dată.

Exemplul 4. Fie $f(x, y)$ o funcție reală, definită și continuă pentru $0 \leqslant x \leqslant 1$, $0 \leqslant y \leqslant 1$. Să considerăm valorile lui $x \in [0, 1]$, pentru care ecuația $f(x, y) = 0$ are o singură soluție $y \in [0, 1]$. Să notăm cu $A(f)$ mulțimea tuturor valorilor lui y care corespund unor astfel de valori ale lui x. $A(f)$ este o mulțime analitică. Reciproc, orice mulțime analitică din $[0, 1]$ este mulțimea $A(f)$ a unei anumite funcții $f(x, y)$ continue pentru $0 \leqslant x \leqslant 1$, $0 \leqslant y \leqslant 1$.

L. Kantorovici numește funcție A (funcție CA) o funcție reală f, pentru care mulțimea $\{x; f(x) > c\}$ este o mulțime analitică (complementară analitică), oricare ar fi c.

Exemplul 5. Fie $f(x, y)$ o funcție mărginită și boreliană. Funcția $g(x) = \sup_y f(x, y)$ este o funcție A, iar funcția $h(x) = \inf_y f(x, y)$ este o funcție CA. Funcția $\limsup_{y \to a} f(x, y)$ este o funcție A, iar $\liminf_{y \to a} f(x, y)$ este o funcție CA.

Exemplul 6. Derivatele parțiale superioare (sau inferioare) ale lui U. Dini ale unei funcții $f(x, y)$ boreliene sînt funcții A (sau CA)

Fie 2^I spațiul tuturor mulțimilor închise conținute în intervalul compact I. Acest spațiu poate fi metrizat cu distanța introdusă de matematicianul român Dimitrie Pompeiu. Fiind date două mulțimi închise $E \subset I$, $F \subset I$, distanța între E și F este, prin definiție, cel mai mare dintre următoarele două numere

$$\sup_{x \in E} \rho(x, F), \sup_{y \in F} \rho(y, E),$$

unde prin $\rho(x, F)$ (respectiv $\rho(y, E)$) am notat marginea inferioară a distanțelor lui x la diversele puncte din F (res-

pectiv marginea inferioară a distanțelor lui y la diversele puncte din E). În spațiul 2^I cu distanța astfel introdusă se pot defini, ca și pe dreaptă, mulțimile închise și deschise. Mulțimile boreliene din 2^I vor fi definite ca mulțimile aparținînd celui mai mic corp care conține mulțimile închise. Prin definiție, o parte a lui 2^I este analitică, dacă ea este imagine continuă a unei părți boreliene.

Cu această pregătire putem trece la

Exemplul 7. În spațiul 2^I, mulțimea avînd drept elemente mulțimile închise nenumărabile din I, este o mulțime analitică neboreliană.

Să considerăm acum spațiul $\mathcal{C}^I$ al funcțiilor reale, continue pe intervalul compact I. Distanța între două elemente x și y din $\mathcal{C}^I$ se definește ca

$$\sup_{t \in I} |x(t) - y(t)|.$$

În spațiul $\mathcal{C}^I$ se definesc, ca și în spațiul 2^I, mai întîi părțile boreliene, apoi cele analitice, apoi complementarele analitice.

Exemplul 8. În spațiul $\mathcal{C}^I$, funcțiile derivabile în fiecare punct din I formează o complementară analitică neboreliană.

Iată acum și un exemplu din teoria funcțiilor de variabilă complexă:

Exemplul 9. Mulțimea valorilor asimptotice ale unei funcții meromorfe în cercul unitate este o mulțime analitică. Reciproc, orice mulțime analitică este mulțimea valorilor asimptotice ale unei anumite funcții meromorfe în cercul unitate.

Vom indica, în cele ce urmează, două exemple de probleme a căror tratare utilizează în mod esențial mulțimile analitice și complementarele lor.

PROBLEMA UNIFORMIZĂRII MULȚIMILOR SAU PROBLEMA FUNCȚIILOR IMPLICITE

Să considerăm mai întîi așa-numita problemă a uniformizării mulțimilor. Fie în planul XOY o mulțime E de puncte. Vom spune că E este uniformă în raport cu axa

OX, dacă orice paralelă la axa OY întîlnește mulțimea E în cel mult un punct.

Vom spune că mulțimea E este uniformizabilă, dacă există o parte uniformă E_1 a lui E, avînd pe OX aceeași proiecție ca și E. E_1 se numește uniformizatoarea lui E.

Este de observat că existența lui E_1 este asigurată de axioma alegerii. Se pune însă problema dacă mulțimea uniformizatoare E_1 poate fi aleasă astfel încît proprietățile ei să fie cît mai apropiate de cele ale lui E. Această problemă este studiată, într-un caz particular, în analiză, la capitolul despre funcții implicite. În adevăr, acolo se dă o funcție continuă $f(x, y)$, prevăzută cu anumite calități, și se uniformizează mulțimea închisă definită de relația $f(x, y) = 0$, cu o mulțime de asemenea închisă.

Ce se întîmplă însă dacă mulțimea E a punctelor de coordonate x, y, pentru care $f(x, y) = 0$, nu mai este închisă? Problema s-a complicat și analiza clasică nu mai este în stare să răspundă la ea. Totuși, problema este deosebit de importantă, atît pentru analiza clasică, cît și pentru analiza funcțională. Este important, de pildă, să se știe dacă E, fiind o mulțime boreliană, poate fi uniformizată cu o mulțime E_1 de asemenea boreliană. Este vizibil că aici este vorba despre o generalizare a problemei funcțiilor implicite.

Nu este greu de văzut că răspunsul la problema de mai sus este negativ. În adevăr, știm că există o mulțime E boreliană în plan (chiar de tip G_δ) a cărei proiecție pe OX este o mulțime E^* neboreliană (deoarece orice mulțime analitică liniară este proiecția unui anume G_δ). Dacă E ar putea fi uniformizată cu o mulțime F_1 boreliana, atunci mulțimea E^* ca imagine biunivocă și continuă (proiecția este o transformare continuă) a lui E_1, ar trebui să fie de asemenea boreliană.

N. Luzin și P.S. Novikov au arătat mai mult, anume că există mulțimi boreliene care nu pot fi uniformizate cu mulțimi boreliene, deși proiecția lor pe OX este boreliană.

Se pune atunci întrebarea, dacă nu cumva orice mulțime boreliană se poate uniformiza cu o mulțime analitică.

Această problemă este importantă, între altele, și pentru faptul că dacă mulțimea punctelor de coordonate $(x, f(x))$

este analitică, atunci se poate arăta că funcția f este măsurabilă. Și aici răspunsul este negativ. S-a putut însă arăta că orice mulțime boreliană din plan se poate uniformiza cu ajutorul unei complementare analitice. În sfîrșit, în ce privește mulțimile analitice, ele nu se pot uniformiza, în general, nici măcar cu complementare analitice. În schimb, după cum a arătat matematicianul japonez Kondo, orice complementară analitică se poate uniformiza cu ajutorul unei complementare analitice.

Menționăm că rezultatele de mai sus asupra uniformizării mulțimilor, ca și cele privitoare la curbele analitice, despre care vom vorbi îndată au fost publicate pentru prima dată de către Luzin, în revista românească „Mathematica" de la Cluj (vol. 4, 1930 și vol. 10, 1935).

PROBLEMA CURBELOR ANALITICE

Un fenomen deosebit de interesant, în ce privește importanța mulțimilor analitice, îl prezintă problema curbelor-analitice. Înțelegem — aici — prin „curbă analitică plană" o mulțime plană analitică, uniformă în raport cu OX. O mulțime boreliană, uniformă în raport cu OX, este o „curbă plană boreliană". S-a putut arăta că dacă o funcție f este boreliană, atunci graficul ei este o curbă boreliană și reciproc. Enunțul acestui rezultat nu conține nici o referire la mulțimi analitice. Cu toate acestea, demonstrația lui nu s-a putut efectua decît cu ajutorul mulțimilor analitice. În fapt, s-a demonstrat că dacă graficul unei funcții este o mulțime analitică, atunci funcția este boreliană, și că dacă funcția este boreliană, atunci și graficul ei este o mulțime boreliană. Problema de a se găsi pentru teorema citată mai sus o altă demonstrație, fără mulțimi analitice, nu a fost rezolvată. W. Sierpiński este de părere că această problemă este prea grea chiar și în cazuri particulare, ca de pildă cel în care graficul funcției este o mulțime boreliană de tip G_δ.

În orice caz, trebuie să reținem că în cadrul mulțimilor plane uniforme noțiunile de mulțime boreliană și

mulțime analitică sînt indiscernabile. Nu același lucru se poate spune despre mulțimile uniforme complementare analitice. Acestea sînt uneori neboreliene.

MULȚIMI PROIECTIVE

Vom da acum noțiunea de mulțime proiectivă și unele proprietăți ale ei.

Mulțimile boreliene sînt, prin definiție, mulțimi proiective de clasă zero. Mulțimile proiective de clasă $2n+1$ sînt imaginile continue ale mulțimilor proiective de clasă $2n$; mulțimile proiective de clasă $2n$ sînt complementarele mulțimilor proiective de clasă $2n$—1. Este vizibil că mulțimile proiective de clasă 1 coincid cu mulțimile analitice, iar cele de clasă 2 coincid cu complementarele analitice. Se poate arăta că o reuniune numărabilă sau o intersecție numărabilă de mulțimi proiective de clasă n este tot o mulțime proiectivă de clasă n. Ca și noțiunea de mulțime analitică sau de complementară analitică, noțiunea de mulțime proiectivă de clasă n este un invariant topologic.

După această introducere în teoria mulțimilor proiective, se ridică în mod firesc unele întrebări: care este justificarea denumirii de „proiective" dată acestor mulțimi? Există, pentru orice număr natural n, o mulțime proiectivă de clasă n care nu este de clasă inferioară lui n?

La prima întrebare răspunsul este simplu. Mulțimile proiective mai pot fi definite ca fiind obținute plecînd de la mulțimi boreliene, prin aplicarea alternativă, de un număr finit de ori, a operațiilor de proiecție și de trecere la complementară. Proiecția trebuie înțeleasă efectuîndu-se dintr-un spațiu euclidian pe un spațiu cu o dimensiune mai puțin. În felul acesta, dacă dorim să obținem mulțimile liniare proiective de ordin n, va trebui să pornim de la mulțimile boreliene dintr-un spațiu euclidian cu un număr suficient de mare de dimensiuni.

Răspunsul la întrebarea a doua este mai complicat, dar mai interesant. Să observăm mai întîi că dat fiind

că mulțimile analitice din spațiul euclidian n-dimensional se obțin ca proiecții ale mulțimilor de tip G_δ din spațiul $n+1$-dimensional, rezultă că putem obține mulțimile proiective liniare de clasă oricît de mare pornind de la mulțimile închise situate într-un spațiu euclidian de dimensiune destul de mare și aplicînd succesiv operațiile de proiecție (ortogonală) pe spațiul cu o dimensiune mai puțin și de trecere la complementară. Este suficient, pentru aceasta, să observăm că orice mulțime de tip F_σ din spațiul euclidian n-dimensional este proiecția ortogonală a unei anumite mulțimi închise din spațiul $n+1$-dimensional.

MULȚIMI UNIVERSALE

Vom introduce acum noțiunea de „mulțime universală" pentru o familie F de mulțimi. Fie F o familie de mulțimi liniare. O mulțime U din plan se numește universală în raport cu familia F, dacă orice mulțime din familia F se poate obține ca intersecție a mulțimii U cu o anumită dreaptă paralelă cu axa ordonatelor. În mod analog se introduce noțiunea de „mulțime universală" pentru spații de dimensiune mai mare.

Să presupunem, pentru moment, că în spațiul euclidian $n+1$-dimensional există o mulțime închisă U universală pentru mulțimile închise în spațiul euclidian cu n-dimensiuni. În acest caz, proiecția mulțimii U în spațiul n-dimensional va fi o mulțime universală pentru mulțimile de tip F_σ din spațiul euclidian n—1-dimensional, iar complementarea acestei proiecții va fi o mulțime G_δ, universală pentru mulțimile G_δ din spațiul n—1-dimensional. Ținînd seama de observația făcută mai sus, proiecția acestei mulțimi G_δ universală va fi o mulțime analitică universală pentru mulțimile analitice din spațiul euclidian cu n—2-dimensiuni. Continuînd tot așa, este lesne de înțeles că vom obține, în cele din urmă, o mulțime plană proiectivă de clasă oricît de mare dorim, universală pentru mulțimile liniare de aceeași clasă proiectivă.

Trebuie numai să avem grijă să pornim cu un n destul de mare.

Să rezumăm: am demonstrat că dacă există mulțimi închise universale, atunci există, pentru orice n, mulțimi proiective de clasă n, universale. Vom demonstra ceva mai jos existența unei mulțimi închise universale. Deocamdată, vom arăta că o mulțime U plană, proiectivă de clasă n, universală, nu poate aparține unei clase proiective inferioare lui n.

UN EXEMPLU SIMPLU DE MULȚIME ANALITICĂ NEBORELIANĂ

Fie mai întîi $n = 1$. Trebuie să arătăm că U nu este boreliană. Pentru aceasta, este suficient să arătăm că $\mathbf{C}U$ (complementara lui U) nu este analitică. În adevăr, să presupunem că $\mathbf{C}U$ este analitică. Fie D o dreaptă din plan, neparalelă cu nici una dintre axele de coordonate. $D \cap \mathbf{C}U$ este analitică. Deci proiecția P a acestei intersecții pe axa ordonatelor este analitică. Deoarece U este universală pentru mulțimile analitice liniare, există a, astfel încît dreapta $x = a$ întîlnește în U o mulțime Π, a cărei proiecție pe axa ordonatelor să fie tocmai P. Fie Q punctul de intersecție al dreptei D cu dreapta $x = a$. Se pot ivi două situații:

1. $Q \notin D \cap U$. În acest caz, proiecția q a lui Q pe $x = 0$ nu aparține lui P, după însăși definiția mulțimii P. Însă, ținînd seamă de definiția mulțimii Π, rezultă că punctul Q nu aparține lui Π, deci nici lui U, ceea ce este contrar ipotezei.

2. $Q \in D \cap U$, deci $Q \in D \cap \mathbf{C}U$, deci $q \in P$. Aceasta înseamnă că $Q \in \Pi$, de unde rezultă că $Q \in U$, adică $Q \in D \cap U$, deoarece Q este un punct al dreptei D. Am ajuns iarăși la o contradicție.

Pentru a putea trage concluzia că U este o mulțime analitică neboreliană, mai trebuie să arătăm că există mulțimi închise universale. Să trecem la această demonstrație.

Fie $I_1, I_2, \ldots I_n, \ldots$ un șir format cu toate intervalele deschise liniare, cu extremități raționale. Este vizibil că orice mulțime liniară este complementara unei anu-

mite reuniuni de astfel de intervale, deci orice mulțime închisă liniară este caracterizată, definită, de un anumit șir de astfel de intervale.

Fie acum un număr irațional x cuprins între 0 și 1 și fie dezvoltarea lui în fracție continuă:

$$x = \cfrac{1}{r_x^1 + \cfrac{1}{r_x^2 + \cfrac{1}{r_x^3 + \dots}}}$$

Să asociem lui x mulțimea închisă F_x dată de complementara mulțimii $I_{r_x^1} \cup I_{r_x^2} \cup I_{r_x^3} \cup \dots$. Reciproc, fiecărei mulțimi liniare închise F îi asociem numărul irațional a cărui dezvoltare în fracție continuă are drept cîturi necomplete indicii de ordine ai intervalelor din șirul $\{I_n\}$, care alcătuiesc complementara lui F. Se stabilește în felul acesta o corespondență biunivocă între mulțimea numerelor iraționale din (0 ,1) și mulțimea mulțimilor închise liniare.

Să definim acum o mulțime plană E în modul următor: punctul de coordonate x, y aparține lui E, dacă x este irațional și cuprins între 0 și 1, iar y aparține mulțimii închise F_x, asociate lui x. Mulțimea E nu este închisă; dar dacă îi adăugăm punctele ei de acumulare, atunci obținem o mulțime U închisă. Vom arăta că U este o mulțime universală pentru mulțimile închise liniare.

Mai întîi să observăm că dacă x_0 este irațional, atunci orice punct din U situat pe dreapta $x = x_0$ aparține obligatoriu lui E. În adevăr, fie $I = (\alpha, \beta)$ un interval din complementara mulțimii închise $E \cap \{x = x_0\}$. Numerele iraționale suficient de apropiate de x au o dezvoltare în fracție continuă, care coincide, pînă la un rang oricît de depărtat dorim, cu dezvoltarea în fracție continuă a lui x_0. Aceasta înseamnă că există $\varepsilon > 0$, astfel încît pentru orice x^* irațional, pentru care $| x - x_0 | < \varepsilon$, intervalul (α, β) este conținut în intersecția $E \cap \{x = x^*\}$. De aici rezultă că intervalul $I = (\alpha, \beta)$ nu poate conține puncte din E'.

Să arătăm acum universalitatea lui U. Fie, pentru aceasta o mulțime F închisă liniară. Fie ξ numărul irațional

asociat lui F prin corespondența biunivocă stabilită mai sus. Este vizibil că dreapta $x = \xi$ întîlnește în mulțimea U exact mulțimea închisă F.

Metoda dezvoltată mai sus, numită metoda mulțimilor universale, este apreciată de obicei ca o metodă efectivă de demonstrație. Este însă evident că această metodă este o variantă a diagonalei lui Cantor. Aceasta a apărut clar în demonstrarea neanaliticității lui $\mathbf{C}U$. Dar poate că acest lucru va apărea mai clar, dacă vom arăta că intersecția dintre U și dreapta D este o mulțime analitică liniară neborelieană (amintim că D este o dreaptă oarecare, neparalelă cu vreuna dintre axele de coordonate). În adevăr, fie Γ această intersecție. Γ este evident analitică. Dacă Γ ar fi borelieană, atunci $D - \Gamma$ ar fi de asemenea borelieană, deci analitică. Fie atunci H proiecția ortogonală a mulțimii $D - \Gamma$ pe axa ordonatelor. H este analitică, fiind imagine continuă a mulțimii analitice $D - \Gamma$. Există atunci o dreaptă $x = \eta$ care întîlnește în U o mulțime analitică H^*, a cărei proiecție pe axa ordonatelor este tocmai H. Fie S punctul de intersecție al lui H^* cu dreapta D. S aparține lui $D - \Gamma$, deoarece se proiectează într-un punct din H. În același timp, S aparține lui H^*, deci lui U, deci lui $\Gamma = D \cap U$. Am ajuns astfel la o contradicție care demonstrează neanaliticitatea mulțimii $D - \Gamma$.

Pentru a se demonstra existența mulțimilor proiective de orice clasă finită, procedăm prin inducție completă, ținînd seamă că dacă o mulțime proiectivă de clasă $2n + 1$ are complementara de aceeași clasă $2n + 1$, atunci ele sînt de clasa $2n$. În felul acesta, este suficient să stabilim existența claselor proiective de rang impar. Aceasta atrage de la sine și existența claselor proiective de rang par.

PROBLEME NEREZOLVATE

Dacă în unele privințe teoria mulțimilor proiective este o generalizare aproape imediată a teoriei mulțimilor analitice, trebuie să spunem că există unele probleme

care, deși la mulțimi analitice și-au găsit o rezolvare relativ simplă, nu au putut fi încă rezolvate pentru mulțimile proiective. Iată cîteva astfel de probleme.

Se știe că mulțimile analitice nenumărabile au puterea continuului. În ce privește însă complementarele analitice și proiecțiile acestor complementare, se știe doar că puterea lor nu poate fi niciodată mai mică decît a continuului și mai mare decît alef unu (alef unu este puterea mulțimii claselor lui Baire; se știe că această mulțime este nenumărabilă). Dar problema puterii acestor mulțimi nu a fost complet rezolvată (în ipoteza neacceptării ipotezei continuului).

În ce privește mulțimile proiective de clasă mai ridicată, trebuie să spunem că despre puterea lor nu se știe încă nimic.

Problema puterii mulțimilor proiective ar fi rezolvată, dacă s-ar putea răspunde afirmativ la următoarea întrebare: o mulțime proiectivă nenumărabilă conține obligatoriu o mulțime perfectă nevidă? Dar această problemă n-a putut fi rezolvată nici măcar pentru complementarele analitice. Ea este apreciată ca una dintre cele mai grele probleme ale matematicii.

Am semnalat că mulțimile analitice și complementarele lor sînt măsurabile Lebesgue. În ce privește însă mulțimile proiective de clasă $n \geqslant 3$, problema măsurabilității lor nu a fost rezolvată.

Natura specială a ultimelor două probleme rezultă din următoarele fapte stabilite de K. Gödel și P.S. Novikov. Există o complementară analitică nenumărabilă, pentru care ipoteza că nu conține nici o mulțime perfectă nevidă nu este în contradicție cu sistemul de axiome al teoriei mulțimilor. Există o mulțime proiectivă pentru care ipoteza nemăsurabilității Lebesgue nu este în contradicție cu sistemul de axiome al teoriei mulțimilor.

Aceste rezultate arată că, indiferent de soluția pe care o vor căpăta problemele semnalate mai sus, este exclusă posibilitatea unor teoreme ca: „orice complementară analitică nenumărabilă conține o mulțime perfectă" sau „orice mulțime proiectivă este măsurabilă Lebesgue".

Se spune că o mulțime E posedă proprietatea lui Baire, dacă este de forma $(G - P) \cup R$, unde G este deschisă, iar P și R sînt de prima categorie. Proprietatea lui Baire este analogul descriptiv al măsurabilității lui Lebesgue. Mulțimile analitice și complementarele lor posedă această proprietate. Nu se cunoaște însă dacă mulțimile proiective de clasă $n \geqslant 3$ posedă sau nu proprietatea lui Baire.

Mai sînt numeroase probleme deschise cu privire la mulțimile proiective, mai cu seamă în legătură cu separabilitatea și cu uniformizarea acestor mulțimi. Dificultatea lor este foarte mare. Unele contribuții esențiale au fost aduse în această problemă de P.S. Novikov.

Rolul mulțimilor proiective în teoria mulțimilor și a funcțiilor apare chiar în studiul unor fenomene relativ simple. Astfel, o mulțime analitică plană nu poate fi uniformizată totdeauna — după cum am semnalat — cu o mulțime analitică și nici măcar cu o complementară analitică. În schimb, uniformizarea unei mulțimi analitice se poate efectua totdeauna cu ajutorul unei proiecții de complementare analitice. Iată acum un exemplu de mulțime proiectivă de clasă 3, care apare în teoria funcțiilor: În spațiul funcțiilor reale, de două variabile reale, continue pe pătratul unitate, funcțiile $f(x, y)$, pentru care există un y, astfel încît derivata parțială $f'_x(x, y)$ există oricare ar fi $x \in [0, 1]$, formează o mulțime proiectivă de clasă 3, care nu este de clasă inferioară lui 3.

O reuniune numărabilă de mulțimi proiective nu mai este, în general, o mulțime proiectivă. Clasificarea mai departe a mulțimilor, luînd ca punct de plecare mulțimile proiective, a fost efectuată de L. Kantorovici și E. Livenson. Ei au studiat, sistematic, operațiile analitice generale introduse de F. Hausdorff și, independent, de A. Kolmogorov. Aceste operații conțin, ca un caz particular, operația prin care Suslin obținea mulțimile analitice ca nuclee ale sistemelor determinante de intervale.

Teoria mulțimilor analitice și proiective a fost extinsă la anumite spații topologice. În cadrul spațiilor metrice complete separabile, teoria lor este o simplă transpunere a teoriei clasice. Încercarea de a face teoria mulțimilor

proiective în spații topologice mai generale decît cele arătate s-a lovit de dificultăți. Ea a fost doar în parte realizată, datorită în primul rînd matematicienilor polonezi și japonezi.

*

Unele dintre cele mai importante contribuții în teoria mulțimilor analitice și proiective au fost publicate în revista „Mathematica" din Cluj. Astfel, articolul lui Luzin *Unele observații asupra curbelor complementare analitice*, publicat în vol. 10 din „Mathematica", în 1934, este reprodus în întregime în noua ediție a cărții lui Luzin, apărută la Moscova în 1953. De asemenea, trebuie să semnalăm că W. Sierpiński, una dintre figurile cele mai proeminente ale matematicii poloneze, care a adus o contribuție fundamentală în teoria mulțimilor analitice și proiective, a publicat o seamă de articole pe această temă în „Mathematica" din Cluj. W. Sierpiński are și meritul de a fi popularizat în rîndul matematicienilor români aceste preocupări, prin lecțiile pe care le-a ținut la Universitatea din Cluj în anii premergători celui de-al doilea război mondial.

Aceste lecții se află editate într-o broșură în limba română.

Florin Vasilescu, în teza sa de doctorat din 1925 asupra funcțiilor multiforme de variabile reale, s-a referit printre primii matematicieni români la mulțimile analitice și proiective. În cartea sa, la sfîrșitul capitolului despre funcții implicite, Luzin semnalizează într-o notă specială teza lui Florian Vasilescu.

Asupra importanței mulțimilor analitice și proiective au insistat, cu diverse ocazii, și alți matematicieni români. Astfel, Simion Stoilow în cartea sa, *Principiile topologice ale teoriei funcțiilor analitice* (în limba franceză), atrage atenția asupra faptului stabilit de Kierst că orice mulțime analitică este mulțimea valorilor asimptotice ale unei anumite funcții meromorfe în cercul unitate.

To the readers

The Mathematical Analysis has been included in the curricula of numerous institutes of higher education since a longtime. Wide sections of intelectuals come into contact with elements of this discipline, but for objective reasons, they have not the possibility to penetrate deeper in its problems and methods. The present book is intended for this large category of readers who, because their type of occupation and lack of time, cannot study systematically and deeply the Mathematical Analysis, but who desire however to get some general view about this science. We address ourselves to everybody who at least once, read a book of Mathematical Analysis of University level. Absorbed by the technical difficulties with which he is faced, the reader of a treatise of Mathematical Analysis, sometimes neglects the very living, developing, side of analysis, of its trends, the problems which give birth to new notions, the syntheses which the analysis operates. But once the technicalities overcomed, the reader may feel the need of a retrospective look at the analysis, his main attention being focused on the origin, significance and evolution of its basic notions, with the view of deciferring its future development.

We would be very glad if the book will be of some help on this road.

Solomon Marcus

Contents

Preface

I

From Eulerian functions to arbitrary functions; from arbitrary functions to computable functions

Evolution of the notion of function until Euler 7
The study of trigonometric series and the generalization of the notion of function 9
The deceptive character of the general notion of function 10
The statistical point of view 14
Continuity phenomena on discontinuous functions 17
Criticism of the classical notions and pleading for the modern notions 19
Theory of distributions, a superior synthesis 22
Degrees of effectiveness 29
What is a normal algorithm? 30
The computable functions as starting point in constructive analysis 35
Turing machines and computable functions 36

II

All types of limits admit a common scheme

Introduction 53
What is a filter? 54
Superior bound of a linear set of points 57
Superior limit of a linear infinite set 58
Superior bound of a function defined on a set in R^n 59
Superior bound of a function in a point 60
Superior or inferior limit of a function in a point 61

Limit of a function in a point 61
Limit of a sequence 62
Superior or inferior limit of a sequence 63
Total variation of a function 65
Darboux integrals of a bounded function 68

III

What is the length of a curve?

Introduction 70
A classical exposition of the notion of the length of the circle 71
Is there a reason to consider inscribed poligonal lines? 72
Difficulties and stumblings in the attempts to define the notion of a curve 74
Curves in analysis 78
What is a rectifiable path? 81
A criterium of rectificability 83
The notion of a rectificable curve 87
Lebesque's criticism concerning the notion „length of a curve" 89
Convergence in distance and convergence in direction 94
Integral representation of the length of a curve 99
Necessity of introduction of a new notion of integral: Lebesque integral 102
The length of a curve, as an inferior semicontinuous functional on the space of curves
The point of view of Fréchet 104

IV

Itinerary in the theory of the integral

Introduction 108
How to determine the area of a figure, more complicated than a polygon 110
From Archimede to Cavalieri 112
Another way to determine the area of a segment of a parable 114
What to make if the function is not monotone 117
Two possible ways 119
When the bases of rectangles diminish 121
How to define the mechanical work of a variable force 121

Similarities and differences between the above tho examples.
The notion of integral 123
Birth of Mathematical Analysis 127
Cauchy moment 132
Riemann moment 133
„Divorce" between primitive and integral 136
Two ideas of Lebesgue 138
Lebesgue measure 148
How extensive is the class of measurable sets? 146
Somethings on Lebesgue measurability of functions 152
A comparison between Riemann integral and Lebesgue integral 156
Reparations after an injustice 168
Infirmities of the Lebesgue integral 175
At the essence of the idea of measure 178
Denjoy integrals and Perron integral 182
From Stieltjes to Riesz. A new perspective in the theory of integral 189

V

In the world of non-borelian sets and functions

Introduction 193
A well inspired mistake 194
Borelian sets and functions 197
Some unusual examples 201
Analytic sets and Suslin operation 204
Lebesgue, involuntary author of the first example of an analytic non-borelian set 206
Sieve of Luzin 210
Analytic sets and irrational numbers 211
How to recognize that an analytic set is borelian? 213
Properties of analytic sets 214
Examples of analytic sets 215
Uniformization of sets and implicit functions 217
Analytic curves 219
Projective sets 220
Universal sets 221
An analytic non-borelian set 222
Unsolved problems 224

Cuprinsul

Către cititori 5

I

De la funcţii euleriene la funcţii arbitrare; de la funcţii arbitrare la funcţii calculabile

Evoluţia noţiunii de funcţie, pînă la Euler 7
Studiul seriilor trigonometrice grăbeşte lărgirea noţiunii de funcţie 9
Caracterul înşelător al noţiunii generale de funcţie 10
Punctul de vedere statistic 14
Fenomene de continuitate la funcţii discontinue 17
O critică a noţiunilor clasice şi o pledoarie pentru noţiunile moderne 19
O sinteză superioară: teoria distribuţiilor 22
Grade de efectivitate 29
Ce este un algoritm normal? 30
Funcţiile calculabile, punct de plecare în analiza constructivă 35
Funcţii calculabile în sensul lui Turing 36

II

Toate tipurile de trecere la limită admit o schemă comună

Introducere 53
Ce este un filtru? 54
Marginea superioară a unei mulţimi liniare de puncte 57
Limita superioară a unei mulţimi liniare, infinite 58
Marginea superioară a unei funcţii definite pe o mulţime din R'' 59
Marginea superioară a unei funcţii într-un punct 60
Limita superioară (respectiv inferioară) a unei funcţii într-un punct 61
Limita unei funcţii într-un punct 61
Limita unui şir 62
Limita superioară (respectiv inferioară) a unui şir 63
Variaţia totală a unei funcţii 65
Integralele lui Darboux ale unei funcţii mărginite 68

III

Ce este lungimea unei curbe?

Introducere 70
O expunere clasică a noţiunii de lungime a cercului 71
De ce linii poligonale înscrise? 72
Dificultăţi şi poticniri în încercările de definire a noţiunii de curbă 74
Noţiunea de curbă în analiză 78
Ce este un drum rectificabil? 81
Un criteriu de rectificabilitate 83
Noţiunea de curbă rectificabilă 87
Critica adusă de Lebesgue teoriei clasice a lungimii 89
Convergenţă în distanţă şi convergenţă în direcţie 94
Reprezentarea integrală a lungimii unei curbe 99
Necesitatea de a se introduce o nouă noţiune de integrală, integrala lui Lebesgue 102
Lungimea unei curbe, ca funcţională inferior semicontinuă în spaţiul curbelor. Punctul de vedere al lui Fréchet 104

IV

Itinerar în teoria integralei

Introducere 108
Cum determinăm aria unei figuri mai complicate? 110
De la Arhimede la Cavalieri 112
Un alt mod de a determina aria unui segment de parabolă 114
Cazul în care funcţia nu mai este monotonă 117
Două căi posibile 119
Ce se întîmplă cînd bazele dreptunghiurilor se micşorează? 121
Ce trebuie să înţelegem prin lucrul mecanic al unei forţe variabile? 121
Asemănări şi deosebiri între cele două exemple de mai sus
Noţiunea de integrală 123
Actul de naştere al analizei matematice 127
Momentul Cauchy 132
Momentul Riemann 133
„Divorţul" dintre primitivă şi integrală 136
Două idei ale lui Lebesgue 138
Măsura Lebesgue 143
Cît de largă este clasa mulţimilor măsurabile Lebesgue? 146
Cîte ceva despre măsurabilitatea Lebesgue a funcţiilor 152
O comparaţie între integrala Riemann şi integrala Lebesgue 156
Reparaţii după o injustiţie 168
Infirmităţi ale integralei Lebsegue 175
La esenţa ideii de măsură 178
Integralele lui Denjoy şi integrala lui Perron 182
De la Stieltjes la Riesz. O nouă perspectivă în teoria integralei 189

V

Ce se întîmplă dincolo de mulțimile și funcțiile boreliene?

Introducere 193
O greșeală bine inspirată 194
Mulțimi și funcții boreliene 197
Cîteva exemple „rare" 201
Mulțimi analitice și operația lui Suslin 204
Lebesgue, autor involuntar al primului exemplu de mulțime analitică neboreliană 208
Sita lui Luzin 210
Mulțimile analitice și mulțimea numerelor iraționale 211
Cum recunoaștem dacă o mulțime analitică este boreliană? 213
Proprietăți ale mulțimilor analitice 214
Exemple de mulțimi analitice 215
Problema uniformizării mulțimilor sau problema funcțiilor implicite 217
Problema curbelor analitice 219
Mulțimi proiective 220
Mulțimi universale 221
Un exemplu simplu de mulțime analitică neboreliană 222
Probleme nerezolvate 224

Prezentare și sumar în limba engleză 229

Redactor resp. de carte : MARIA BORICEAN
Tehnoredactor : VICTORIA STĂNCULESCU

Dat la cules 27.05.1967. Bun de tipar 20.10.1967. Tiraj 8500+140 ex. Hîrtie scris II A 63 g/m². Format 90×100/32. Coli editoriale 11,83. Coli tipar 7,50 A. 6706. Indici de clasificare zecimală : pentru bibliotecile mari 517, pentru bibliotecile mici 51.

Tiparul executat sub comanda nr. 439 la
Întreprinderea Poligrafică „13 Decembrie 1918",
Str. Grigore Alexandrescu nr. 93—95,
București — Republica Socialistă România

ERATĂ

Pag.	*Rînd*	*În loc de:*	*Se va citi*
67	1 sus 2 sus 5 jos	$\triangle$	$\triangle$
133	13 sus	natural	real

Lucrarea "Noțiuni de analiză matematică"
de S. Marcus

Lei 5,75

editura științifică

Lucrarea prezintă, fără a pune accent pe aspectele tehnice, apariția principalelor noțiuni ale analizei matematice, semnificația lor și evoluția pînă la stadiul actual. Astfel, sînt expuse noțiunile de funcție, limită, lungimea unei curbe, aria unei suprafețe, integrală, precum și unele noțiuni de teoria mulțimilor și topologie. Se evită calculele și demonstrațiile mai grele, accentul fiind pus pe evoluția ideilor. Lucrarea se adresează elevilor din ultimele clase de liceu, studenților, profesorilor de liceu și, în general, celor interesați în aprofundarea problematicii și metodelor analizei matematice, ea oferind o retrospectivă deosebită, și originală asupra dezvoltării analizei și a tendințelor ce o animă.